Invariance Principles for Homogeneous Sums: Universality of Gaussian Wiener Chaos

Ivan Nourdin, Giovanni Peccati, Gesine Reinert

Invariance Principles for Homogeneous Sums: Universality of Gaussian Wiener Chaos

by Ivan Nourdin*, Giovanni Peccati† and Gesine Reinert‡

Université Paris VI, Université Paris Ouest and Oxford University

Abstract: We compute explicit bounds in the normal and chi-square approximations of multilinear homogenous sums (of arbitrary order) of general centered independent random variables with unit variance. Our techniques combine an invariance principle by Mossel, O'Donnell and Oleszkiewicz with a refinement of some recent results by Nourdin and Peccati, about the approximation of laws of random variables belonging to a fixed (Gaussian) Wiener chaos. In particular, we show that chaotic random variables enjoy the following form of *universality*: (a) the normal and chi-square approximations of any homogenous sum can be completely characterized and assessed by first switching to its Wiener chaos counterpart, and (b) the simple upper bounds and convergence criteria available on the Wiener chaos extend almost verbatim to the class of homogeneous sums. These results partially rely on the notion of "low influences" for functions defined on product spaces, and provide a generalization of central and non-central limit theorems proved by Nourdin, Nualart and Peccati. They also imply a further drastic simplification of the method of moments and cumulants – as applied to the proof of probabilistic limit theorems – and yield substantial generalizations, new proofs and new insights into some classic findings by de Jong and Rotar'. Our tools involve the use of Malliavin calculus, and of both the Stein's method and the Lindeberg invariance principle for probabilistic approximations.

Key words: Central Limit Theorems; Chaos; Homogeneous Sums; Lindeberg Principle; Malliavin Calculus; Chi-square Limit Theorems; Stein's Method; Universality; Wiener Chaos.

2000 Mathematics Subject Classification: 60F05; 60F17; 60G15; 60H07.

*Laboratoire de Probabilités et Modèles Aléatoires, Université Pierre et Marie Curie (Paris VI), Boîte courrier 188, 4 place Jussieu, 75252 Paris Cedex 05, France. Email: `ivan.nourdin@upmc.fr`

†Equipe Modal'X, Université Paris Ouest – Nanterre la Défense, 200 Avenue de la République, 92000 Nanterre, France, and LSTA, Université Paris VI, France. Email: `giovanni.peccati@gmail.com`

‡Department of Statistics, University of Oxford, 1 South Parks Road, Oxford OX1 3TG, UK. Email `reinert@stats.ox.ac.uk`

Contents

1 Introduction

1.1 Overview

The aim of this paper is to study and characterize the normal and chi-square approximations of the laws of *multilinear homogeneous sums* involving general independent random variables. We shall perform this task by implicitly combining three probabilistic techniques, namely: (i) the *Lindeberg invariance principle* (in the version developed by Mossel et al. [20] and Rotar' [40], see also Mossel [19]), (ii) the *Stein's method* for the normal and chi-square approximations (see e.g. [3, 36, 41, 42]), and (iii) the *Malliavin calculus of variations* on a Gaussian space (see e.g. [16] and [27]). Our analysis will reveal that the Gaussian Wiener chaos (see Section 2 below for precise definitions) enjoys the following properties: (**a**) the normal and chi-square approximations of any multilinear homogenous sum are completely characterized and assessed by those of its Wiener chaos counterpart, and (**b**) the strikingly simple upper bounds and convergence criteria available on the Wiener chaos (see [21, 22, 23, 24, 26, 29]) extend almost verbatim to the class of homogeneous sums. Our findings partially rely on the notion of "low influences" (see again [20]) for real-valued functions defined on product spaces. As indicated by the title, we regard the two properties (**a**) and (**b**) as an instance of the *universality phenomenon*, according to which most information about large random systems (such as the "distance to Gaussian" of non-linear functionals of large samples of independent random variables) does not depend on the particular distribution of the components. Other recent examples of the universality phenomenon appear in the already quoted paper [20], as well as in the Tao-Vu proof of the circular law for random matrices, as detailed in [45] (see also the Appendix to [45] by Krishnapur). Observe that in Section 7 we will prove analogous results for the multivariate normal approximation of *vectors* of homogenous sums of possibly different orders.

1.2 The approach

In what follows, every random object is defined on a suitable (common) probability space $(\Omega, \mathscr{F}, P)$. The symbol E denotes expectation with respect to P. We start by giving a precise definition of the main objects of our study.

Definition 1.1 (Multilinear homogeneous sums) Fix some integers $N, d \geqslant 2$ and write $[N] = \{1, ..., N\}$. Let $\mathbf{X} = \{X_i : i \geqslant 1\}$ be a collection of centered independent random variables, and let $f : [N]^d \to \mathbb{R}$ be a *symmetric function vanishing on diagonals* (that is, $f(i_1, ..., i_d) = 0$ whenever there exist $k \neq j$ such that $i_k = i_j$). The random variable

$$Q_d(N,f,\mathbf{X}) = Q_d(\mathbf{X}) = \sum_{1\leqslant i_1,\dots,i_d\leqslant N} f(i_1,\dots,i_d)X_{i_1}\cdots X_{i_d} \tag{1.1}$$

$$= d! \sum_{\{i_1,\dots,i_d\}\subset[N]^d} f(i_1,\dots,i_d)X_{i_1}\cdots X_{i_d} \tag{1.2}$$

$$= d! \sum_{1\leqslant i_1<\dots<i_d\leqslant N} f(i_1,\dots,i_d)X_{i_1}\cdots X_{i_d} \tag{1.3}$$

is called the *multilinear homogeneous sum*, of order d, based on f and on the first N elements of $\mathbf{X}$. As in the first equality of (1.1), and when there is no risk of confusion, we will drop the dependence on N and f in order to simplify the notation. Plainly, $E[Q_d(\mathbf{X})] = 0$ and also, if $E(X_i^2) = 1$ for every i, then

$$E[Q_d(\mathbf{X})^2] = d!\|f\|_d^2, \tag{1.4}$$

where, here and for the rest of the paper, we use the notation

$$\|f\|_d^2 = \sum_{1\leqslant i_1,\dots,i_d\leqslant N} f^2(i_1,\dots,i_d).$$

In the following, we will systematically use the expression "homogeneous sum" instead of "multilinear homogeneous sum".

Objects such as the RHS of (1.1) are sometimes called "polynomial chaoses", and play a central role in several branches of probability theory and stochastic analysis. For instance:

– In the case $d = 2$, homogeneous sums are indeed standard *quadratic forms* with no diagonal terms. Limit theorems for quadratic forms have been the object of an intense study since 30 years. An exhaustive discussion about the best available results on the approximation of the laws of quadratic forms can be found in Götze and Tikhomirov [9].

– By setting $X_i = H_{q_i}(G_i)$, where $\{G_i\}$ is an independent and identically distributed (i.i.d.) centered standard Gaussian sequence and H_q is the Hermite polynomial of degree q, one obtains the basic building blocks of the so-called *Gaussian Wiener chaos* (see Section 2), whose properties are at the core of the Malliavin calculus (see e.g. [11, 27]) and of the white noise analysis (see e.g. [30]).

– By setting $X_i = J_{q_i}(P_i)$, where J_q indicates the Charlier polynomial of degree q and $\{P_i\}$ are the values of a centered Poisson measure over disjoint sets, one obtains a class of random elements generating the *Poisson Wiener chaos* (see e.g. [31, 32, 35, 38]).

– By setting $\mathbf{X}$ to be a *Rademacher sequence* (that is, the X_i's are i.i.d. and such that $P(X_1 = 1) = P(X_1 = -1) = 1/2$) then $Q_d(\mathbf{X})$ are the constitutive elements of the so-called *Walsh chaos*, playing a central role e.g. in harmonic analysis (see [1]), probability on Banach spaces (see [13]), quantum probability (see [18, Ch. 2]), social choice theory and theoretical computer sciences (see [20], and the references therein, for an overview of results in these last two areas).

Despite the almost ubiquitous nature of homogeneous sums, results concerning the normal approximation of quantities such as (1.1) in the non-quadratic case (i.e., when $d \geqslant 3$) are surprisingly scarce: indeed, to our knowledge, the only general statements in this respect are contained in the paper [6] and the monograph [5], both by P. de Jong (as discussed below, and in a different direction, general criteria allowing to assess the proximity of the laws of homogenous sums based on different independent sequences are obtained in [20, 39, 40]).

Remark 1.2 We stress that (1.1) can be viewed as a sum of *locally dependent* random variables, with index set $\mathcal{J} = \{1, \ldots, n\}^d$. Note moreover that the techniques usually adopted in the normal approximation of sums of dependent variables are not appropriate here. For instance, one cannot adopt the approach developed by Chen and Shao in [2]: indeed, in our framework the neighborhoods of dependence

$$A_{\mathbf{i}} = \{\mathbf{j} = (j_1, \ldots, j_d) \in \mathcal{J} : |\mathbf{i} \cap \mathbf{j}| \neq 0\}, \quad \mathbf{i} = (i_1, \ldots, i_d) \in \mathcal{J},$$

verify the relation $\cup_{\mathbf{j} \in A_{\mathbf{i}}} A_{\mathbf{j}} = \mathcal{J}$ and are consequently "too large".

In this paper, we are interested in controlling objects of the type

$$d_{\mathscr{H}}\{Q_d(\mathbf{X})\,;\,Z\},$$

where: (i) $Q_d(\mathbf{X})$ is defined in (1.1), (ii) Z is either a standard Gaussian $\mathscr{N}(0,1)$ or a centered chi-square random variable, and (iii) the distance $d_{\mathscr{H}}\{F;G\}$, between the laws of two random variables F, G, is given by

$$d_{\mathscr{H}}\{F\,;\,G\} = \sup\big\{\big|E[h(F)] - E[h(G)]\big| : h \in \mathscr{H}\big\}, \tag{1.5}$$

(E denotes expectation) with $\mathscr{H}$ some suitable class of real-valued functions.

Even with some uniform control on the components of $\mathbf{X}$, the problem of directly and generally assessing $d_{\mathscr{H}}\{Q_d(\mathbf{X})\,;\,Z\}$ looks very arduous. Indeed, any estimate comparing the laws of $Q_d(\mathbf{X})$ and Z capriciously depends on the kernel f, and on the way in which the analytic structure of f interacts with the specific "shape" of the distribution of the random variables X_i. One revealing picture of this situation appears if one tries to evaluate the moments of $Q_d(\mathbf{X})$ and to compare them with those of Z: see e.g. [33, 43] for a discussion of some associated combinatorial structures. In the specific case where Z is Gaussian, one should also observe that $Q_d(\mathbf{X})$ is a *completely degenerate U-statistic*, as $E[Q_d(X_1; x_2, \ldots, x_d)] = 0$ for all $x_2, \ldots, x_d$, so that the standard results for the normal approximation of U-statistics do not apply.

The main point developed in this paper is that one can successfully overcome these difficulties by implementing the following strategy: first (**I**) measure the distance

$$d_{\mathscr{H}}\{Q_d(\mathbf{X})\,;\,Q_d(\mathbf{G})\},$$

between the law of $Q_d(\mathbf{X})$ and the law of the random variable $Q_d(\mathbf{G})$, obtained by replacing $\mathbf{X}$ with a centered standard i.i.d. Gaussian sequence $\mathbf{G} = \{G_i : i \geqslant 1\}$, then (**II**) assess the distance $d_{\mathscr{H}}\{Q_d(\mathbf{G})\,;\,Z\}$, and finally (**III**) use the triangle inequality in order to write

$$d_{\mathscr{H}}\{Q_d(\mathbf{X})\,;\,Z\} \leqslant d_{\mathscr{H}}\{Q_d(\mathbf{X})\,;\,Q_d(\mathbf{G})\} + d_{\mathscr{H}}\{Q_d(\mathbf{G})\,;\,Z\}. \tag{1.6}$$

We will see in the subsequent sections that the power of this approach resides in the following two facts.

– The distance evoked at Point (**I**) can be effectively controlled by means of the techniques developed in [20], where the authors have produced a general theory, based on a generalization of the so-called *Lindeberg invariance principle* (see e.g. [46]), allowing to estimate the distance between multilinear homogeneous sums constructed from different sequences of independent random variables. A full discussion of this point is presented in Section 4 below. In Theorem 4.1 we shall observe that, under the assumptions that $E(X_i^2) = 1$ and that the moments $E(|X_i|^3)$ are uniformly bounded by some constant $\beta > 0$ (recall that the X_i's are centered), one can deduce from [20] that (provided that the elements of $\mathscr{H}$ are sufficiently smooth),

$$d_{\mathscr{H}}\{Q_d(\mathbf{X})\,;\,Q_d(\mathbf{G})\} \leqslant C \times \sqrt{\max_{1\leqslant i\leqslant N} \sum_{\{i_2,\ldots,i_d\}\in[N]^{d-1}} f^2(i,i_2,\ldots,i_d)}, \tag{1.7}$$

where C is a constant depending on d, β and on the class $\mathscr{H}$. The quantity

$$\mathrm{Inf}_i(f) = \sum_{\{i_2,\ldots,i_d\}\in[N]^{d-1}} f^2(i,i_2,\ldots,i_d) = \frac{1}{(d-1)!} \sum_{1\leqslant i_2,\ldots,i_d\leqslant N} f^2(i,i_2,\ldots,i_d) \tag{1.8}$$

is called the *influence* of the variable i, and roughly quantifies the contribution of X_i to the overall configuration of the homogenous sum $Q_d(\mathbf{X})$. Influence indices already appear (under a different name) in the papers by Rotar' [39, 40].

– The random variable $Q_d(\mathbf{G})$ is an element of the dth *Wiener chaos* associated with $\mathbf{G}$ (see Section 2 for definitions). As such, the distance between $Q_d(\mathbf{G})$ and Z (in both the normal and the chi-square cases) can be assessed by means of the results appearing in [21, 22, 23, 24, 26, 28, 29, 34], which are in turn based on a powerful interaction between standard Gaussian analysis, Stein's method, and the Malliavin calculus on variations. As an example, Theorem 3.1 of Section 3 proves that the above quoted results imply that, if $Q_d(\mathbf{G})$ has variance one and Z is standard Gaussian, then

$$d_{\mathscr{H}}\{Q_d(\mathbf{G})\,;\,Z\} \leqslant C\sqrt{|E[Q_d(\mathbf{G})^4] - E(Z^4)|} = C\sqrt{|E[Q_d(\mathbf{G})^4] - 3|} \tag{1.9}$$

where $C > 0$ is some finite constant depending only on $\mathscr{H}$ and d. Moreover, if the elements of $\mathbf{X}$ have bounded fourth moments, then the RHS of (1.9) can be directly estimated by means of the fourth moment of $Q_d(\mathbf{X})$ and some maxima over the influence functions appearing in (1.7) (see Section 5 for a full discussion of this point).

1.3 Universality

Bounds such as (1.6), (1.7) and (1.9) only partially account for the term "universality" appearing in the title of the present paper. Our techniques allow indeed to prove the following statement, involving vectors of homogeneous sums of possibly different orders, see Theorem 7.5.

Theorem 1.3 (Universality of Wiener chaos) *Let $\mathbf{G} = \{G_i : i \geqslant 1\}$ be a standard centered i.i.d. Gaussian sequence, and fix integers $m \geqslant 1$ and $d_1, ..., d_m \geqslant 2$. For every $j = 1, ..., m$, let $\{(N_n^{(j)}, f_n^{(j)}) : n \geqslant 1\}$ be a sequence such that $\{N_n^{(j)} : n \geqslant 1\}$ is a sequence of integers going to infinity, and each function $f_n^{(j)} : [N_n^{(j)}]^{d_j} \to \mathbb{R}$ is symmetric and vanishes on diagonals. Define $Q_{d_j}(N_n^{(j)}, f_n^{(j)}, \mathbf{G})$, $n \geqslant 1$, according to (1.1). Assume that, for every $j = 1, ...m$, the following sequence of variances is bounded:*

$$E[Q_{d_j}(N_n^{(j)}, f_n^{(j)}, \mathbf{G})^2], \quad n \geqslant 1. \tag{1.10}$$

Let V be a $m \times m$ non-negative symmetric matrix, and let $\mathscr{N}_m(0, V)$ indicate a centered Gaussian vector with covariance matrix V. Then, as $n \to \infty$, the following two conditions are equivalent.

(1) *The vector $\{Q_{d_j}(N_n^{(j)}, f_n^{(j)}, \mathbf{G}) : j = 1, ..., m\}$ converges in law to $\mathscr{N}_m(0, V)$.*

(2) *For every sequence $\mathbf{X} = \{X_i : i \geqslant 1\}$ of independent centered random variables, with unit variance and such that $\sup_i E|X_i|^3 < \infty$, the law of the vector $\{Q_{d_j}(N_n^{(j)}, f_n^{(j)}, \mathbf{X}) : j = 1, ..., m\}$ converges to the law of $\mathscr{N}_m(0, V)$ in the Kolmogorov distance.*

Remark 1.4 1. Given random vectors $F = (F_1, ..., F_m)$ and $H = (H_1, ..., H_m)$, $m \geqslant 1$, the *Kolmogorov distance* between the law of F and the law of H is defined as

$$d_{Kol}(F, H) = \sup_{(z_1,...,z_m)\in\mathbb{R}^m} |P(F_1 \leqslant z_1, ..., F_m \leqslant z_m) - P(H_1 \leqslant z_1, ..., H_m \leqslant z_m)|. \tag{1.11}$$

Recall that the topology induced by d_{Kol} on the class of all probability measures on $\mathbb{R}^m$ is strictly stronger than the topology of convergence in distribution.

2. Note that, in the statement of Theorem 1.3, we do not require that the matrix V is *positively definite*, and we do *not* introduce any assumption on the asymptotic behavior of influence indices.

3. Due to the matching moments up to second order, one has that

$$E[Q_{d_i}(N_n^{(i)}, f_n^{(i)}, \mathbf{G}) \times Q_{d_j}(N_n^{(j)}, f_n^{(j)}, \mathbf{G})] = E[Q_{d_i}(N_n^{(i)}, f_n^{(i)}, \mathbf{X}) \times Q_{d_j}(N_n^{(j)}, f_n^{(j)}, \mathbf{X})]$$

for every $i, j = 1, ..., m$ and every sequence $\mathbf{X}$ as at Point (2) of Theorem 1.3.

Plainly, the crucial implication in the statement of Theorem 1.3 is (1) $\Rightarrow$ (2), which basically ensures that any statement concerning the asymptotic normality of (vectors of) general homogeneous sums can be proved by simply focussing on the elements of a Gaussian Wiener chaos. Since Central Limit Theorems (CLTs) on Wiener chaos are by now completely characterized (thanks to the results proved in [22, 24, 28, 29, 34]), this fact represents a clear methodological breakthrough. As explained later in the paper, and up to the restriction on the third moments appearing at Point (2), we regard Theorem 1.3 as the first exact equivalent – for homogeneous sums – of the usual CLT for linear functionals of i.i.d. sequences. The proof of Theorem 1.3 is completely achieved in Section 7.

Remark 1.5

1. The use of the techniques developed in [20] makes it unavoidable to require a uniform bound on the third moments of $\mathbf{X}$. However, one advantage of using Lindeberg-type arguments is that we obtain convergence in the Kolmogorov distance, as well as explicit upper bounds on the rates of convergence. We will see below (see Theorem 1.12 for a precise statement) that in the one-dimensional case one can simply require a bound on the moments of order $2+\epsilon$, for some $\epsilon > 0$. Moreover, still in the one-dimensional case and when the sequence $\mathbf{X}$ is i.i.d., one can alternatively deduce convergence in distribution from a result by Rotar' [40, Proposition 1], for which the existence of moments of order greater than 2 is not required.

2. In this paper we will apply the techniques of [20] in order to evaluate quantities of the type $|E[Q_d(\mathbf{X})^m] - E[Q_d(\mathbf{Y})^m]|$, where $m \geqslant 1$ is an integer and $\mathbf{X}$ and $\mathbf{Y}$ are two sequences of independent random variables. From a purely methodological standpoint, we believe that this technique may replace the (sometimes fastidious) use of "diagram formulae" and associated combinatorial devices (see [33, 43]). See Section 4 for a full discussion of this point, especially Lemma 4.4.

1.4 The role of contractions

The universality principle stated in Theorem 1.3 is based on [20], as well as on general characterizations of (possibly multidimensional) CLTs on a fixed Wiener chaos. Results of this kind have been first proved in [29] (for the one-dimensional case) and [34] (for the multidimensional case), and make an important use of the notion of "contraction" of a given deterministic kernel (see also [28]). When studying homogeneous sums, one is naturally led to deal with contractions defined on discrete sets of the type $[N]^d$, $N \geqslant 1$. As already pointed out in [25], these objects enjoy a number of useful combinatorial properties, related for instance to the notion of *fractional dimension* or *fractional Cartesian* *product* (see Blei [1]). In this section, we shall briefly explore these issues, in particular by pointing out that discrete contractions are indeed the key element in the proof of Theorem 1.3. More general statements, as well as complete proofs, are given in Section 3.

Definition 1.6 Fix $d, N \geqslant 2$. Let $f : [N]^d \to \mathbb{R}$ be a symmetric function vanishing of diagonals. For every $r = 0, ..., d$, the *contraction* $f \star_r f$ is the function on $[N]^{2d-2r}$ given by

$$\begin{aligned} & f \star_r f(j_1, ..., j_{2d-2r}) \\ & = \sum_{1 \leqslant a_1, ..., a_r \leqslant N} f(a_1, ..., a_r, j_1, ..., j_{d-r}) f(a_1, ..., a_r, j_{d-r+1}, ..., j_{2d-2r}). \end{aligned} \tag{1.12}$$

Observe that $f \star_r f$ is not necessarily symmetric and not necessarily vanishes on diagonals. The symmetrization of $f \star_r f$ is written $f \widetilde{\star}_r f$.

The following result, whose proof will be achieved in Section 7, as a special case of Theorem 7.5, is based on the findings of [29] and [34].

Proposition 1.7 (CLT for chaotic sums) *Let the assumptions and notations of Theorem 1.3 prevail, and suppose moreover that, for every $i, j = 1, ..., m$ (as $n \to \infty$)*

$$E[Q_{d_i}(N_n^{(i)}, f_n^{(i)}, \mathbf{G}) \times Q_{d_j}(N_n^{(j)}, f_n^{(j)}, \mathbf{G})] \to V(i, j), \tag{1.13}$$

where V is a non-negative symmetric matrix. Then, the following three conditions are equivalent, as $n \to \infty$.

(1) *The vector $\{Q_{d_j}(N_n^{(j)}, f_n^{(j)}, \mathbf{G}) : j = 1, ..., m\}$ converges in law to a centered Gaussian vector with covariance matrix V.*

(2) *For every $j = 1, ..., m$, $E[Q_{d_j}(N_n^{(j)}, f_n^{(j)}, \mathbf{G})^4] \to 3V(i, i)^2$.*

(3) *For every $j = 1, ..., m$ and every $r = 1, ..., d_j - 1$, $\|f_n^{(j)} \star_r f_n^{(j)}\|_{2d_j - 2r} \to 0$.*

Remark 1.8 Strictly speaking, the results of [34] only deal with the case where V is positive definite. The needed general result will be obtained in Section 7 by means of Malliavin calculus.

Let us now briefly sketch the proof of the implication (1) $\Rightarrow$ (2) in Theorem 1.3. Suppose that the sequence in (1.10) is bounded and that Point (1) in Theorem 1.3 is verified. Then, by uniform integrability (use Proposition 2.7), the convergence (1.13) is satisfied and, according to Proposition 1.7, we have $\|f_n^{(j)} \star_{d_j-1} f_n^{(j)}\|_2 \to 0$. The crucial remark is now that

$$
\begin{aligned}
\|f_n^{(j)} \star_{d_j-1} f_n^{(j)}\|_2^2 &= \sum_{1\leqslant i,k\leqslant N_n^{(j)}} \left[\sum_{1\leqslant i_2,\dots,i_{d_j}\leqslant N_n^{(j)}} f_n^{(j)}(i,i_2,\dots,i_{d_j}) f_n^{(j)}(k,i_2,\dots,i_{d_j}) \right]^2 \\
&\geqslant \sum_{i=1}^{N_n^{(j)}} \left[\sum_{1\leqslant i_2,\dots,i_{d_j}\leqslant N_n^{(j)}} f_n^{(j)}(i,i_2,\dots,i_{d_j})^2 \right]^2 \\
&\geqslant \max_{i=1,\dots,N_n^{(j)}} \left[\sum_{1\leqslant i_2,\dots,i_{d_j}\leqslant N_n^{(j)}} f_n^{(j)}(i,i_2,\dots,i_{d_j})^2 \right]^2 \\
&= \left[(d_j-1)! \max_{1\leqslant i\leqslant N_n^{(j)}} \mathrm{Inf}_i(f_n^{(j)}) \right]^2 \qquad (1.14)
\end{aligned}
$$

(recall formula (1.8)), from which one immediately obtains that, as $n \to \infty$,

$$\max_{1\leqslant i\leqslant N_n^{(j)}} \mathrm{Inf}_i(f_n^{(j)}) \to 0, \quad \text{for every } j = 1,\dots,m. \qquad (1.15)$$

The proof of Theorem 1.3 is concluded by using Theorem 7.1, which is a multi-dimensional version of the findings of [20]. Indeed, this result will imply that, if (1.15) is verified then, for every sequence $\mathbf{X}$ as in Point (2) of Theorem 1.3, the distance between the law of $\{Q_{d_j}(N_n^{(j)}, f_n^{(j)}, \mathbf{G}) : j = 1,\dots,m\}$ and the law of $\{Q_{d_j}(N_n^{(j)}, f_n^{(j)}, \mathbf{X}) : j = 1,\dots,m\}$ necessarily tends to zero, and therefore the two sequences must converge in distribution to the same limit.

As proved in [21], contractions play an equally important role in the chi-square approximation of the laws of elements of a fixed chaos of even order. Recall that a random variable Z_ν has a *centered chi-square distribution* with $\nu \geqslant 1$ degrees of freedom (noted $Z_\nu \sim \chi^2(\nu)$) if $Z_\nu \overset{\text{Law}}{=} \sum_{i=1}^{\nu} \left(G_i^2 - 1\right)$, where $(G_1,\dots,G_\nu)$ is a vector of i.i.d. centered Gaussian random variables with unit variance. Note that

$$E(Z_\nu^2) = 2\nu, \quad E(Z_\nu^3) = 8\nu, \quad \text{and} \quad E(Z_\nu^4) = 12\nu^2 + 48\nu. \qquad (1.16)$$

The following result is proved in [21].

Theorem 1.9 (Chi-square limit theorem for chaotic sums) *Let $\mathbf{G} = \{G_i : i \geqslant 1\}$ be a standard centered i.i.d. Gaussian sequence, and fix an even integer $d \geqslant 2$. Let $\{N_n, f_n : n \geqslant 1\}$ be a sequence such that $\{N_n : n \geqslant 1\}$ is a sequence of integers going to infinity, and each $f_n : [N_n]^d \to \mathbb{R}$ is symmetric and vanishes on diagonals. Define $Q_d(N_n, f_n, \mathbf{G})$, $n \geqslant 1$, according to (1.1), and assume that, as $n \to \infty$, $E[Q_d(N_n, f_n, \mathbf{G})^2] \to 2\nu$. Then, as $n \to \infty$, the following three conditions are equivalent.*

(1) $Q_d(N_n, f_n, \mathbf{G}) \overset{\text{Law}}{\to} Z_\nu \sim \chi^2(\nu)$.

(2) $E[Q_d(N_n, f_n, \mathbf{G})^4] - 12E[Q_d(N_n, f_n, \mathbf{G})^3] \to E[Z_\nu^4] - 12E[Z_\nu^3] = 12\nu^2 - 48\nu$.

(3) $\|f_n \widetilde{\star}_{d/2} f_n - c_d \times f_n\|_d \to 0$ *and* $\|f_n \star_r f_n\|_{2d-2r} = 0$ *for every* $r = 1, ..., d-1$ *such that* $r \neq d/2$, *where*

$$c_d := \frac{1}{(d/2)!\binom{d-1}{d/2-1}^2} = \frac{4}{(d/2)!\binom{d}{d/2}^2}. \tag{1.17}$$

A neat application of Theorem 1.9 is sketched in the next section.

1.5 Example: revisiting de Jong criterion

To further clarify the previous discussion (and to motivate the reader), we provide an illustration of how one can use our results in order to refine a remarkable result by de Jong, originally proved in [6].

Theorem 1.10 (See [6]) *Let* $\mathbf{X} = \{X_i : i \geqslant 1\}$ *be a sequence of independent centered random variables such that* $E(X_i^2) = 1$ *and* $E(X_i^4) < \infty$ *for every* i. *Fix* $d \geqslant 2$, *and let* $\{N_n, f_n : n \geqslant 1\}$ *be a sequence such that* $\{N_n : n \geqslant 1\}$ *is a sequence of integers going to infinity, and each* $f_n : [N_n]^d \to \mathbb{R}$ *is symmetric and vanishes on diagonals. Define* $Q_d(n, \mathbf{X}) = Q_d(N_n, f_n, \mathbf{X})$, $n \geqslant 1$, *according to (1.1). Assume that* $E[Q_d(n, \mathbf{X})^2] = 1$ *for all* n. *If, as* $n \to \infty$,

(i) $E[Q_d(n, \mathbf{X})^4] \to 3$, *and*

(ii) $\max_{1 \leqslant i \leqslant N_n} \sum_{1 \leqslant i_2, ..., i_d \leqslant N_n} f_n^2(i, i_2, ..., i_d) = (d-1)! \max_{1 \leqslant i \leqslant N_n} \mathrm{Inf}_i(f_n) \to 0$,

then $Q_d(n, \mathbf{X})$ *converges in law to a standard Gaussian random variable* $Z \sim \mathscr{N}(0, 1)$.

Remark 1.11 As shown in [6], the conclusion of Theorem 1.10 extends to homogeneous sums that are not necessarily multilinear. Also, the fact that $E(X_i^4) < \infty$ does not appear in [6] (however, an inspection of the proof of Theorem 1.10 shows that this assumption is indeed necessary).

In the original proof given in [6], Assumption (i) in Theorem 1.10 appears as a convenient (and mysterious) way of reexpressing the asymptotic "lack of interaction" between products of the type $X_{i_1} \cdots X_{i_d}$, whereas Assumption (ii) plays the role of a usual Lindeberg-type assumption. In the present paper, under the slightly stronger assumption that $\beta := \sup_i E(X_i^4) < \infty$, we will be able to produce bounds neatly indicating the exact roles of both Assumption (i) and Assumption (ii). To see this, define $d_{\mathscr{H}}$ according to (1.5), and set $\mathscr{H}$ to be the class of thrice differentiable functions whose first three derivatives are bounded by some finite constant $B > 0$. In Section 5, in the proof of Theorem 5.1 we will show that there exist universal, explicit, finite constants $C_1, C_2, C_3 > 0$, depending uniquely on β, d and B, such that (writing $\mathbf{G}$ for an i.i.d. centered standard Gaussian sequence)

$$\begin{aligned} d_{\mathscr{H}}\{Q_d(n,\mathbf{X});Q_d(n,\mathbf{G})\} &\leqslant C_1 \times \left(\max_{1\leqslant i\leqslant N_n} \mathrm{Inf}_i(f_n) + \sqrt{\max_{1\leqslant i\leqslant N_n} \mathrm{Inf}_i(f_n)} \right) & (1.18) \\ d_{\mathscr{H}}\{Q_d(n,\mathbf{G});Z\} &\leqslant C_2 \times \sqrt{|E[Q_d(n,\mathbf{G})^4]-3|} & (1.19) \\ |E[Q_d(n,\mathbf{X})^4]-E[Q_d(n,\mathbf{G})^4]| &\leqslant C_3 \times \left(\max_{1\leqslant i\leqslant N_n} \mathrm{Inf}_i(f_n) + \sqrt{\max_{1\leqslant i\leqslant N_n} \mathrm{Inf}_i(f_n)} \right). & (1.20) \end{aligned}$$

In particular, the estimates (1.18) and (1.20) show that Assumption (ii) in Theorem 1.10 ensures that both the laws and the fourth moments of $Q_d(n,\mathbf{X})$ and $Q_d(n,\mathbf{G})$ are asymptotically close: this fact, combined with Assumption (i) implies that the LHS of (1.19) converges to zero. This gives an alternate proof of Theorem 1.10 in the case of uniformly bounded fourth moments.

Also, by combining the universality principle stated in Theorem 1.3 with (1.19) (or, alternatively, with Proposition 1.7 in the case $m = 1$) and with [40, Proposition 1], one obtains the following "universal version" of de Jong's criterion.

Theorem 1.12 *Let $\mathbf{G} = \{X_i : i \geqslant 1\}$ be a centered i.i.d. Gaussian sequence with unit variance. Fix $d \geqslant 2$, and let $\{N_n, f_n : n \geqslant 1\}$ be a sequence such that $\{N_n : n \geqslant 1\}$ is a sequence of integers going to infinity, and each $f_n : [N_n]^d \to \mathbb{R}$ is symmetric and vanishes on diagonals. Define $Q_d(n,\mathbf{G}) = Q_d(N_n, f_n, \mathbf{G})$, $n \geqslant 1$, according to (1.1). Assume that $E[Q_d(n,\mathbf{G})^2] \to 1$ as $n \to \infty$. Then, the following four properties are equivalent as $n \to \infty$.*

(1) *The sequence $Q_d(n,\mathbf{G})$ converges in law to $Z \sim \mathscr{N}(0,1)$.*

(2) $E[Q_d(n,\mathbf{G})^4] \to 3$.

(3) *For every sequence $\mathbf{X} = \{X_i : i \geqslant 1\}$ of independent centered random variables with unit variance and such that $\sup_i E|X_i|^{2+\epsilon} < \infty$ for some $\epsilon > 0$, the sequence $Q_d(n,\mathbf{X})$ converges in law to $Z \sim \mathscr{N}(0,1)$ in the Kolmogorov distance.*

(4) *For every sequence $\mathbf{X} = \{X_i : i \geqslant 1\}$ of independent and identically distributed centered random variables with unit variance, the sequence $Q_d(n,\mathbf{X})$ converges in law to $Z \sim \mathscr{N}(0,1)$ (not necessarily in the Kolmogorov distance).*

Remark 1.13 1. Note that at point (4) of the above statement we *do not* require the existence of moments of order greater than 2. We will see that the equivalence between (1) and (4) is partly a consequence of Rotar's results [40].

2. Theorem 1.12 is a particular case of Theorem 1.3, and can be seen as refinement of de Jong's Theorem 1.10, in the sense that: (i) since several combinatorial devices

are at hand (see e.g. [33]), it is in general easier to evaluate moments of multilinear forms of Gaussian sequences than of general sequences, and (ii) when the $\{X_i\}$ are not identically distributed, we only need existence (and uniform boundedness) of the moments of order $2+\epsilon$.

In Section 7, we will generalize the content of this section to multivariate Gaussian approximations. By using Proposition 1.9 and [40, Proposition 1], one can also obtain the following universal chi-square limit result.

Theorem 1.14 *We let the notation of Theorem 1.12 prevail, except that we now assume that $d \geqslant 2$ is an even integer and $E[Q_d(n,\mathbf{G})^2] \to 2\nu$, where $\nu \geqslant 1$ is an integer. Then, the following conditions are equivalent as $n \to \infty$.*

(1) *The sequence $Q_d(n,\mathbf{G})$ converges in law to $Z_\nu \sim \chi^2(\nu)$.*

(2) $E[Q_d(n,\mathbf{G})^4] - 12E[Q_d(n,\mathbf{G})^3] \to E(Z_\nu^4) - 12E(Z_\nu^3) = 12\nu^2 - 48\nu$.

(3) *For every sequence $\mathbf{X} = \{X_i : i \geqslant 1\}$ of independent centered random variables with unit variance and such that $\sup_i E|X_i|^{2+\epsilon} < \infty$ for some $\epsilon > 0$, the sequence $Q_d(n,\mathbf{X})$ converges in law to Z_ν.*

(4) *For every sequence $\mathbf{X} = \{X_i : i \geqslant 1\}$ of independent and identically distributed centered random variables with unit variance, the sequence $Q_d(n,\mathbf{X})$ converges in law to Z_ν.*

1.6 Two counterexamples

1.6.1 There is no universality for sums of order one

One striking feature of Theorem 1.3 and Theorem 1.12 is that *they do not have any equivalent for sums of order $d=1$*. To see this, consider an array of real numbers $\{f_n(i) : 1 \leqslant i \leqslant n\}$ such that $\sum_{i=1}^n f_n^2(i) = 1$. Let $\mathbf{G} = \{G_i : i \geqslant 1\}$ and $\mathbf{X} = \{X_i : i \geqslant 1\}$ be, respectively, a centered i.i.d. Gaussian sequence with unit variance, and a sequence of independent random variables with zero mean and unit variance. Then, $Q_1(n,\mathbf{G}) := \sum_{i=1}^n f_n(i)G_i \sim \mathscr{N}(0,1)$ for every n, but it is in general *not true* that $Q_1(n,\mathbf{X}) := \sum_{i=1}^n f_n(i)X_i$ converges in law to a Gaussian random variable. An example of this situation is obtained by taking X_1 to be non-Gaussian, $f_n(1) = 1$ and $f_n(j) = 0$ for $j > 1$. As it is well-known, to ensure that $Q_1(n,\mathbf{X})$ has a Gaussian limit one customarily adds the Lindeberg-type requirement that $\max_{1\leqslant i\leqslant n} |f_n(i)| \to 0$ (see e.g. [12, Th. 4.12]). A closer inspection indicates that the fact that no Lindeberg conditions are required in Theorem 1.3 and Theorem 1.12 is due to the implication (1) $\Rightarrow$ (3) in Proposition 1.7, as well as to the inequality (1.14).

1.6.2 Walsh chaos is not universal

We shall now prove that one cannot replace the Gaussian sequence $\mathbf{G}$ with a Rademacher one in the statements of Theorem 1.3 and Theorem 1.12. Let $\mathbf{X} = \{X_i : i \geqslant 1\}$ be an i.i.d. Rademacher sequence, and fix $d \geqslant 2$. For every $N \geqslant d$, consider the homogeneous sum

$$Q_d(N, \mathbf{X}) = X_1 X_2 \cdots X_{d-1} \sum_{i=d}^{N} \frac{X_i}{\sqrt{N-d+1}} \tag{1.21}$$

It is easily seen that each $Q_d(N, \mathbf{X})$ can be written in the form (1.1), for some symmetric $f = f_N$ vanishing on diagonals and such that $d!\|f_N\|_d^2 = 1$. Since $X_1 X_2 \cdots X_{d-1}$ is a random sign independent of $\{X_i : i \geqslant d\}$, a simple application of the Central Limit Theorem yields that, as $N \to \infty$, $Q_d(N, \mathbf{X}) \overset{\text{Law}}{\to} \mathscr{N}(0,1)$. On the other hand, a direct computation reveals that, for every $N \geqslant d$, $\max_{1 \leqslant i \leqslant N} \mathrm{Inf}_i(f_N) = (1/d!)^2$. This shows that homogenous sums associated with Rademacher sequences *are not* universal. For instance, by replacing $\mathbf{X}$ in (1.21) with a i.i.d. standard Gaussian sequence $\mathbf{G} = \{G_i : i \geqslant 1\}$, one sees that, for every $N \geqslant 2$,

$$Q_d(N, \mathbf{G}) = G_1 G_2 \cdots G_{d-1} \sum_{i=d}^{N} \frac{G_i}{\sqrt{N-d+1}} \overset{\text{Law}}{=} G_1 \cdots G_d. \tag{1.22}$$

Since (for $d \geqslant 2$) the random variable $G_1 \cdots G_d$ is not Gaussian, this yields that the sequence $Q_d(N, \mathbf{G})$, $N \geqslant 1$, does not verify a Central Limit Theorem.

1.7 Three caveats on generality

In order to enhance the readability of the forthcoming material, we decided not to state some of our findings in full generality. In particular:

(i) It will be clear later on that the results of this paper easily extend to the case of *infinite* homogeneous sums (obtained by putting $N = +\infty$ in (1.1)). This requires, however, a somewhat heavier notation, as well as some distracting digressions about convergence.

(ii) Our findings do not hinge at all on the fact that $\mathbb{N}$ is an ordered set: it follows that our results exactly apply to homogeneous sums of random variables indexed by a general finite set.

(iii) Our results on chi-square approximations could be slightly modified in order to accommodate approximations by arbitrary centered Gamma laws. To do this, one can use the results about Gamma approximations proved in [21, 22].

2 Wiener chaos

In this section, we briefly introduce the notion of (Gaussian) *Wiener chaos*, and point out some of its crucial properties. The reader is referred to [27, Ch. 1] or [11, Ch. 2] for any unexplained definition or result.

2.1 Definitions

Let $\mathbf{G} = \{G_i : i \geqslant 1\}$ be a sequence of i.i.d. centered Gaussian random variables with unit variance.

Definition 2.1 (Hermite polynomials and Wiener chaos) 1. The sequence of Hermite polynomials $\{H_q : q \geqslant 0\}$ is defined as follows: $H_0 = 1$, and, for $q \geqslant 1$,

$$H_q(x) = (-1)^q e^{\frac{x^2}{2}} \frac{d^q}{dx^q} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}.$$

Recall that the class $\{(q!)^{-1/2} H_q : q \geqslant 0\}$ is an orthonormal basis of

$$L^2(\mathbb{R}, (2\pi)^{-1/2} e^{-x^2/2} dx).$$

2. A *multi-index* $q = \{q_i : i \geqslant 1\}$ is a sequence of nonnegative integers such that $q_i \neq 0$ only for a finite number of indices i. We also write Λ to indicate the class of all multi-indices, and use the notation $|q| = \sum_{i \geqslant 1} q_i$, for every $q \in \Lambda$.

3. For every $d \geqslant 0$, the dth *Wiener chaos* associated with $\mathbf{G}$ is defined as follows: $C_0 = \mathbb{R}$, and, for $d \geqslant 1$, C_d is the $L^2(P)$-closed vector space generated by random variables of the type

$$\Phi(q) = \prod_{i=1}^{\infty} H_{q_i}(G_i), \quad q \in \Lambda \text{ and } |q| = d. \tag{2.23}$$

Example 2.2 (i) The first Wiener chaos C_1 is the Gaussian space generated by $\mathbf{G}$, that is, $F \in C_1$ if and only if

$$F = \sum_{i=1}^{\infty} \lambda_i G_i, \text{ for some sequence } \{\lambda_i : i \geqslant 1\} \in \ell^2.$$

(ii) Fix $d, N \geqslant 2$ and let $f : [N]^d \to \mathbb{R}$ be symmetric and vanishing on diagonals. Then, the d-homogeneous sum

$$Q_d(\mathbf{G}) = d! \sum_{\{i_1,\ldots,i_d\} \subset [N]^d} f(i_1, \ldots, i_d) G_{i_1} \cdots G_{i_d} = \sum_{1 \leqslant i_1,\ldots,i_d \leqslant N} f(i_1, \ldots, i_d) G_{i_1} \cdots G_{i_d} \tag{2.24}$$

is an element of C_d.

It is easily seen that two random variables belonging to Wiener chaoses of different orders are orthogonal in $L^2(P)$. Moreover, since linear combinations of polynomials are dense in $L^2(P,\sigma(\mathbf{G}))$, one has that $L^2(P,\sigma(\mathbf{G})) = \bigoplus_{d\geqslant 0} C_d$, that is, any square integrable functional of $\mathbf{G}$ can be written as an infinite sum, converging in L^2 and such that the dth summand is an element of C_d. This representation result is known as the *Wiener-Itô chaotic decomposition* of $L^2(P,\sigma(\mathbf{G}))$.

It is often useful to encode the properties of random variables in the spaces C_d by using increasing tensor powers of Hilbert spaces (see e.g. [11, Appendix E] for a collection of useful facts about tensor products). To do this, introduce a (arbitrary) real separable Hilbert space $\mathfrak{H}$ and, for $d \geqslant 2$, denote by $\mathfrak{H}^{\otimes d}$ (resp. $\mathfrak{H}^{\odot d}$) the dth tensor power (resp. symmetric tensor power) of $\mathfrak{H}$; write moreover $\mathfrak{H}^{\otimes 0} = \mathfrak{H}^{\odot 0} = \mathbb{R}$ and $\mathfrak{H}^{\otimes 1} = \mathfrak{H}^{\odot 1} = \mathfrak{H}$. Let $\{e_j : j \geqslant 1\}$ be an orthonormal basis of $\mathfrak{H}$. With every multi-index $q \in \Lambda$, we associate the tensor $e(q) \in \mathfrak{H}^{\otimes |q|}$ given by

$$e(q) = e_{i_1}^{\otimes q_{i_1}} \otimes \cdots \otimes e_{i_k}^{\otimes q_{i_k}},$$

where $\{q_{i_1}, ..., q_{i_k}\}$ are the non-zero elements of q. We also denote by $\tilde{e}(q) \in \mathfrak{H}^{\odot |q|}$ the canonical symmetrization of $e(q)$. It is well-known that, for every $d \geqslant 2$, the collection $\{\tilde{e}(q) : q \in \Lambda,\, |q| = d\}$ defines a complete orthogonal system in $\mathfrak{H}^{\odot d}$. For every $d \geqslant 1$ and every $h \in \mathfrak{H}^{\odot d}$ with the form $h = \sum_{q\in\Lambda,\,|q|=d} c_q \tilde{e}(q)$, we define

$$I_d(h) = \sum_{q\in\Lambda,\,|q|=d} c_q \Phi(q), \tag{2.25}$$

where $\Phi(q)$ is given in (2.23). The following result (see e.g. [27, p. 8] for a proof) characterizes I_d as an isomorphism.

Proposition 2.3 *For every $d \geqslant 1$, the mapping $I_d : \mathfrak{H}^{\odot d} \to C_d$ (as defined in (2.25)) is onto, and provides an isomorphism between C_d and the Hilbert space $\mathfrak{H}^{\odot d}$, endowed with the norm $\sqrt{d!}\|\cdot\|_{\mathfrak{H}^{\otimes d}}$. In particular, for every $h, h' \in \mathfrak{H}^{\odot d}$,*

$$E[I_d(h)I_d(h')] = d!\langle h, h'\rangle_{\mathfrak{H}^{\otimes d}}.$$

As proved e.g. in [27, Section 1.1.2], if $\mathfrak{H} = L^2(A, \mathscr{A}, \mu)$, with μ σ-finite and non-atomic, then the operators I_d are indeed (multiple) Wiener-Itô integrals.

Example 2.4 (i) By definition, $G_i = I_1(e_i)$, for every $i \geqslant 1$.

(ii) The random variable $Q_d(\mathbf{G})$ defined in (2.24) is such that

$$Q_d(\mathbf{G}) = I_d(h), \quad \text{where} \quad h = d! \sum_{\{i_1,...,i_d\}\subset[N]^d} f(i_1, ..., i_d) e_{i_1} \otimes \cdots \otimes e_{i_d} \in \mathfrak{H}^{\odot d}. \tag{2.26}$$

2.2 Multiplication formulae and hypercontractivity

The notion of "contraction" is the key to prove the general bounds stated in the forthcoming Section 3.

Definition 2.5 (Contractions) Let $\{e_i : i \geqslant 1\}$ be a complete orthonormal system in $\mathfrak{H}$. Given $f \in \mathfrak{H}^{\odot p}$ and $g \in \mathfrak{H}^{\odot q}$, for every $r = 0, \ldots, p \wedge q$, the rth contraction of f and g is the element of $\mathfrak{H}^{\otimes(p+q-2r)}$ defined as

$$f \otimes_r g = \sum_{i_1,\ldots,i_r=1}^{\infty} \langle f, e_{i_1} \otimes \ldots \otimes e_{i_r} \rangle_{\mathfrak{H}^{\otimes r}} \otimes \langle g, e_{i_1} \otimes \ldots \otimes e_{i_r} \rangle_{\mathfrak{H}^{\otimes r}}. \tag{2.27}$$

Plainly, $f \otimes_0 g = f \otimes g$ equals the tensor product of f and g while, for $p = q$, $f \otimes_p g = \langle f, g \rangle_{\mathfrak{H}^{\otimes p}}$. Note that, in general (and except for trivial cases), the contraction $f \otimes_r g$ is *not* a symmetric element of $\mathfrak{H}^{\otimes(p+q-2r)}$. The canonical symmetrization of $f \otimes_r g$ is written $f \widetilde{\otimes}_r g$.

Contractions appear in multiplication formulae like the following one (see [27, Prop. 1.1.3] for a proof).

Proposition 2.6 (Multiplication formulae) *If $f \in \mathfrak{H}^{\odot p}$ and $g \in \mathfrak{H}^{\odot q}$, then*

$$I_p(f) I_q(g) = \sum_{r=0}^{p \wedge q} r! \binom{p}{r} \binom{q}{r} I_{p+q-2r}(f \widetilde{\otimes}_r g). \tag{2.28}$$

Note that the previous statement implies that multiple integrals admit finite moments of every order. The next result (whose proof is given e.g. in [11, Th. 5.10]) establishes a more precise property, namely that random variables living in a finite sum of Wiener chaoses are hypercontractive.

Proposition 2.7 (Hypercontractivity) *Let $d \geqslant 1$ be a finite integer and assume that $F \in \bigoplus_{k=0}^{d} C_k$. Fix reals $2 \leqslant p \leqslant q < \infty$. Then,*

$$E[|F|^q]^{\frac{1}{q}} \leqslant (q-1)^{\frac{d}{2}} E[|F|^p]^{\frac{1}{p}}. \tag{2.29}$$

We will use hypercontractivity in order to deduce some of the bounds stated in the next section.

3 Normal and chi-square approximation on Wiener chaos

Starting from this section, and for the rest of the paper, we adopt the following notation for distances between laws of real-valued random variables.

– The symbol $d_{TV}(F,G)$ indicates the *total variation distance* between the law of F and G, obtained from (1.5) by taking $\mathscr{H}$ equal to the class of all indicators of the Borel subsets of $\mathbb{R}$.

– The symbol $d_W(F,G)$ denotes the *Wasserstein distance*, obtained from (1.5) by choosing $\mathscr{H}$ as the class of all Lipschitz functions with Lipschitz constant less or equal to 1.

– The symbol $d_{BW}(F,G)$ stands for the *bounded Wasserstein distance* (or *Fortet-Mourier distance*), deduced from (1.5) by choosing $\mathscr{H}$ as the class of all Lipschitz functions that are bounded by 1, and with Lipschitz constant less than or equal to 1.

While $d_{Kol}(F,G) \leqslant d_{TV}(F,G)$ and $d_{BW}(F,G) \leqslant d_W(F,G)$, in general $d_{TV}(F,G)$ and $d_W(F,G)$ are not comparable, see [8] for an overview.

In what follows, we consider as given an i.i.d. centered standard Gaussian sequence $\mathbf{G} = \{G_i : i \geqslant 1\}$, and we shall adopt the Wiener chaos notation introduced in Section 2.

3.1 Bounds on the normal approximation

In the recent series of papers [22, 23, 26], it has been shown that one can effectively combine Malliavin calculus with Stein's method, in order to evaluate the distance between the law of an element of a fixed Wiener chaos, say F, and a standard Gaussian distribution. In this section, we state several refinements of these results, by showing in particular that all the relevant bounds can be expressed in terms of the fourth moment of F. The proofs involve the use of Malliavin calculus and are deferred to Section 8.

Theorem 3.1 (Fourth moment bounds) *Fix $d \geqslant 2$. Let $F = I_d(h)$, $h \in \mathfrak{H}^{\odot d}$, be an element of the dth Gaussian Wiener chaos C_d such that $E(F^2) = 1$, let $Z \sim \mathscr{N}(0,1)$, and write*

$$
\begin{aligned}
T_1(F) &:= \sqrt{d^2 \sum_{r=1}^{d-1} (r-1)!^2 \binom{d-1}{r-1}^4 (2d-2r)! \left\| h \widetilde{\otimes}_r h \right\|^2_{\mathfrak{H}^{\otimes 2(d-r)}}} \\
T_2(F) &:= \sqrt{\frac{d-1}{3d} |E(F^4) - 3|}.
\end{aligned}
$$

We have

$$T_1(F) \leqslant T_2(F). \tag{3.30}$$

Moreover, the following bounds hold:

1.

$$d_{TV}(F,Z) \leqslant 2T_1(F). \tag{3.31}$$

2.

$$d_{BW}(F,Z) \leqslant d_W(F,Z) \leqslant T_1(F). \tag{3.32}$$

3. *Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a thrice differentiable function such that $\|\varphi'''\|_\infty < \infty$. Write*

$$\begin{aligned} K_* &= K_*\big(|\varphi'(0)|, |\varphi''(0)|, \|\varphi'''\|_\infty\big) && (3.33)\\ &= \max\left\{2\,|\varphi''(0)| + \frac{\|\varphi'''\|_\infty}{3}\frac{2\sqrt{2}}{\sqrt{\pi}}; 2\,|\varphi'(0)| + \frac{1}{3}\,\|\varphi'''\|_\infty\right\} \\ C_* &= C_*\big(d, |\varphi'(0)|, |\varphi''(0)|, \|\varphi'''\|_\infty\big) = \big(3 + (2\sqrt{2})^d\big) \times K_*. && (3.34) \end{aligned}$$

Then, one has that

$$|E[\varphi(F)] - E[\varphi(Z)]| \leqslant C_* \times T_1(F).$$

Proof. See Section 8.3. ∎

Remark 3.2 If $E(F) = 0$ and F has a finite fourth moment, then the quantity $\kappa_4(F) = E(F^4) - 3E(F^2)^2$ is known as the *fourth cumulant* of F. See e.g. [43] for a general introduction to cumulants. One can also prove (see e.g. [29]) that, if F is a non-zero element of the dth Wiener chaos of a given Gaussian sequence ($d \geqslant 2$), then $\kappa_4(F) > 0$.

3.2 Central limit theorems

Now fix $d \geqslant 2$, and consider a sequence of random variables of the type $F_n = I_d(h_n)$, $n \geqslant 1$, such that, as $n \to \infty$, $E(F_n^2) = d!\|h_n\|^2_{\mathfrak{H}^{\otimes d}} \to 1$. In [29] it is proved that the following double implication holds: as $n \to \infty$,

$$\left\|h_n \widetilde{\otimes}_r h_n\right\|_{\mathfrak{H}^{\otimes 2(d-r)}} \to 0,\ \forall r = 1, ..., d-1 \Leftrightarrow \|h_n \otimes_r h_n\|_{\mathfrak{H}^{\otimes 2(d-r)}} \to 0,\ \forall r = 1, ..., d-1. \tag{3.35}$$

Theorem 3.1, combined with (3.35), allows therefore to recover the following characterization of CLTs on Wiener chaos. It has been first proved (by completely different methods) in [29].

Theorem 3.3 (See [28, 29]) *Fix $d \geqslant 2$, and let $F_n = I_d(h_n)$, $n \geqslant 1$ be a sequence in the dth Wiener chaos of* $\mathbf{G}$. *Assume that $\lim_{n\to\infty} E(F_n^2) = 1$. Then, the following conditions are equivalent, as $n \to \infty$.*

(1) *F_n converges in law to $Z \sim \mathscr{N}(0,1)$.*

(2) *$E(F_n^4) \to E(Z^4) = 3$.*

(3) *For every $r = 1, ..., d-1$, $\|h_n \otimes_r h_n\|_{\mathfrak{H}^{\otimes 2(d-r)}} \to 0$.*

Proof. Since $\sup_n E(F_n^2) < \infty$, one deduces from Proposition 2.7 that, for every $M > 2$, $\sup_n E|F_n|^M < \infty$. By uniform integrability it follows that, if (1) is in order, then necessarily $E(F_n^4) \to E(Z^4) = 3$. The rest of the proof is a consequence of the bounds (3.30)–(3.31) and of the fact that the topology induced by the total variation distance (on the class of all probability measures on $\mathbb{R}$) is stronger than the topology of convergence in distribution. ■

The following (elementary) result is one of the staples of the present paper. It relates the Hilbert space contractions of formula (2.27) with the discrete contraction operators defined in (1.12). We state it in a form which is also useful for the chi-square approximation of the forthcoming Section 3.3.

Lemma 3.4 *Fix $d \geqslant 2$, and suppose that $h \in \mathfrak{H}^{\odot d}$ is given by (2.26), with $f : [N]^d \to \mathbb{R}$ symmetric and vanishing on diagonals. Then, for $r = 1, ..., d-1$,*

$$\|h \otimes_r h\|_{\mathfrak{H}^{\otimes(2d-2r)}} = \|f \star_r f\|_{2d-2r}, \tag{3.36}$$

where we have used the notation (1.12). Also, if d is even, then for every $\alpha_1, \alpha_2 \in \mathbb{R}$

$$\|\alpha_1(h \otimes_{d/2} h) + \alpha_2 h\|_{\mathfrak{H}^{\otimes d}} = \|\alpha_1(f \star_{d/2} f) + \alpha_2 f\|_d. \tag{3.37}$$

Proof. Fix $r = 1, ..., d-1$. Using (2.26) and the fact that $\{e_j : j \geqslant 1\}$ is an orthonormal basis of $\mathfrak{H}$, one infers that

$$\begin{aligned} h \otimes_r h &= \sum_{1\leqslant i_1,...,i_d\leqslant N} \sum_{1\leqslant j_1,...,j_d\leqslant N} f(i_1,...,i_d) f(j_1,...,j_d)[e_{i_1} \otimes \cdots \otimes e_{i_d}] \otimes_r [e_{j_1} \otimes \cdots \otimes e_{j_d}] \\ &= \sum_{1\leqslant a_1,...,a_r\leqslant N} \sum_{1\leqslant k_1,...,k_{2d-2r}\leqslant N} f(a_1,...a_r,k_1,...,k_{d-r}) \times \\ &\qquad\qquad \times f(a_1,...,a_r,k_{d-r+1},...,k_{2d-2r}) e_{k_1} \otimes \cdots \otimes e_{k_{2d-2r}} \\ &= \sum_{1\leqslant k_1,...,k_{2d-2r}\leqslant 1} f \star_r f(k_1,...,k_{2d-2r}) e_{k_1} \otimes \cdots \otimes e_{k_{2d-2r}}. \end{aligned} \tag{3.38}$$

Since the set $\{e_{k_1} \otimes \cdots \otimes e_{k_{2d-2r}} : k_1, ..., k_{2d-2r} \geqslant 1,$ mutually distinct$\}$ is an orthonormal basis of $\mathfrak{H}^{\otimes(2d-2r)}$, one deduces immediately (3.36). The proof of (3.37) is analogous. ■

Remark 3.5 Theorem 3.3 and formula (3.36) yield immediately a proof of Proposition 1.7 in the case $m = 1$.

3.3 Bounds on the chi-square approximation

As demonstrated in [21, 22], the combination of Malliavin calculus and Stein's method also allows to estimate the distance between the law of an element F of a fixed Wiener chaos and a (centered) chi-square distribution $\chi^2(\nu)$ with ν degrees of freedom. Analogously to the previous section for Gaussian approximations, we now state a number of refinements of the results proved in [21, 22]. In particular, we will show that all the relevant bounds can be expressed in terms of a specific linear combination of the third and fourth moments of F. Once again, the proofs are deferred to Section 8.

Theorem 3.6 (Third and fourth moment bounds) *Fix an even integer $d \geqslant 2$ as well as an integer $\nu \geqslant 1$. Let $F = I_d(h)$ be an element of the dth Gaussian chaos C_d such that $E(F^2) = 2\nu$, let $Z_\nu \sim \chi^2(\nu)$, and write*

$$\begin{aligned} T_3(F) &:= \Bigg[4d! \left\| h - \frac{d!^2}{4(d/2)!^3} h \widetilde{\otimes}_{d/2} h \right\|^2_{\mathfrak{H}^{\otimes d}} \\ &\quad + d^2 \sum_{\substack{r=1,\dots,d-1 \\ r \neq d/2}} (r-1)!^2 \binom{d-1}{r-1}^4 (2d-2r)! \left\| h \widetilde{\otimes}_r h \right\|^2_{\mathfrak{H}^{\otimes 2(d-r)}} \Bigg]^{\frac{1}{2}} \\ T_4(F) &:= \sqrt{\frac{d-1}{3d} |E(F^4) - 12E(F^3) - 12\nu^2 + 48\nu|}. \end{aligned}$$

Then,

$$T_3(F) \leqslant T_4(F) \tag{3.39}$$

and

$$d_{BW}(F, Z_\nu) \leqslant \max\left\{ \sqrt{\frac{2\pi}{\nu}}, \frac{1}{\nu} + \frac{2}{\nu^2} \right\} T_3(F). \tag{3.40}$$

Remark 3.7 Observe that $\frac{d!^2}{4(d/2)!^3} = \frac{1}{c_d}$, where c_d is given in (1.17).

Proof. See Section 8.4. ■

3.4 Chi-square limit theorems

Now fix an even integer $d \geqslant 2$, and consider a sequence of random variables of the type $F_n = I_d(h_n)$, $n \geqslant 1$, such that, as $n \to \infty$, $E(F_n^2) = d! \|h_n\|^2_{\mathfrak{H}^{\otimes d}} \to 2\nu$. In [21] it is proved that the following double implication holds: as $n \to \infty$,

$$\begin{aligned} & \left\| h_n \widetilde{\otimes}_r h_n \right\|_{\mathfrak{H}^{\otimes 2(d-r)}} \to 0, \ \forall r = 1, ..., d-1, \ r \neq d/2 \\ \Longleftrightarrow \quad & \left\| h_n \otimes_r h_n \right\|_{\mathfrak{H}^{\otimes 2(d-r)}} \to 0, \ \forall r = 1, ..., d-1, \ r \neq d/2. \end{aligned} \tag{3.41}$$

Theorem 3.6, combined with (3.41), allows therefore to recover the following characterization of chi-square limit theorems on Wiener chaos. Note that this is a special case of a 'non-central limit theorem'; one usually calls 'non-central limit theorem' any result involving convergence in law to a non-Gaussian distribution.

Theorem 3.8 (See [21]) *Fix an even integer $d \geqslant 2$, and let $F_n = I_d(h_n)$, $n \geqslant 1$ be a sequence in the dth Wiener chaos of* **G**. *Assume that $\lim_{n\to\infty} E(F_n^2) = 2\nu$. Then, the following conditions are equivalent, as $n \to \infty$.*

(1) *F_n converges in law to a $Z_\nu \sim \chi^2(\nu)$.*

(2) $E(F_n^4) - 12E(F_n^3) \to E(Z_\nu^4) - 12E(Z_\nu^3) = 12\nu^2 - 48\nu$.

(3) $\|h_n \widetilde{\otimes}_{d/2} h_n - c_d \times h_n\|_{\mathfrak{H}^{\otimes d}} \to 0$ *(for* c_d *as given in (1.17)) and, for every* $r = 1, ..., d-1$ *such that* $r \neq d/2$, $\|h_n \otimes_r h_n\|_{\mathfrak{H}^{\otimes 2(d-r)}} \to 0$.

Proof. Since $\sup_n E(F_n^2) < \infty$, one deduces from Proposition 2.7 that, for every $M > 2$, $\sup_n E|F_n|^M < \infty$. By uniform integrability it follows that, if (1) holds, then necessarily $E(F_n^4) - 12E(F_n^3) \to E(Z_\nu^4) - 12E(Z_\nu^3) = 12\nu^2 - 48\nu$. The rest of the proof is a consequence of the bound (3.39) and of the fact that the topology induced by the bounded Wasserstein distance (on the class of all probability measures on $\mathbb{R}$) is equivalent to the topology of convergence in distribution. ∎

Remark 3.9 By using (3.37) in the case $\alpha_1 = 1$ and $\alpha_2 = -c_d$, one sees that Theorem 3.8 yields an immediate proof of Proposition 1.9.

4 Low influences and proximity of homogeneous sums

We now turn to some remarkable invariance principles by Rotar' [40] and Mossel, O' Donnell and Oleszkiewicz [20]. As already discussed, the results proved in [40] yield sufficient conditions in order to have that the laws of homogeneous sums (or, more generally, polynomial forms) that are built from two different sequences of independent random variables are asymptotically close, whereas in [20] one can find explicit upper bounds on the distance between these laws. Since in this paper we adopt the perspective of deducing general convergence results from limit theorems on a Gaussian space, we will state the results of [40] and [20] in a slightly less general form, namely by assuming that one of the sequences is i.i.d. Gaussian. See also Davydov and Rotar' [4], and the references therein, for some general characterizations of the asymptotic proximity of probability distributions.

Theorem 4.1 (See [20]) *Let* $\mathbf{X} = \{X_i,\ i \geqslant 1\}$ *be a collection of centered independent random variables with unit variance, and let* $\mathbf{G} = \{G_i,\ i \geqslant 1\}$ *be a collection of standard centered i.i.d. Gaussian random variables. Fix* $d \geqslant 1$, *and let* $\{N_n, f_n : n \geqslant 1\}$ *be a sequence such that* $\{N_n : n \geqslant 1\}$ *is a sequence of integers going to infinity, and each* $f_n : [N_n]^d \to \mathbb{R}$ *is symmetric and vanishes on diagonals. Define* $Q_d(N_n, f_n, \mathbf{X})$ *and* $Q_d(N_n, f_n, \mathbf{G})$ *according to (1.1).*

1. *If* $\sup_{i\geqslant 1} E[|X_i|^{2+\epsilon}] < \infty$ *for some* $\epsilon > 0$ *and if* $\max_{1\leqslant i\leqslant N_n} \mathrm{Inf}_i(f_n) \to 0$ *as* $n \to \infty$ *(with* $\mathrm{Inf}_i(f_n)$ *given by (1.8)) then we have*

$$\sup_{z\in\mathbb{R}} \left|P\big[Q_d(N_n, f_n, \mathbf{X}) \leqslant z\big] - P\big[Q_d(N_n, f_n, \mathbf{G}) \leqslant z\big]\right| \to 0 \quad \text{as } n \to \infty. \quad (4.42)$$

2. *If the random variables* X_i *are identically distributed and if* $\max_{1\leqslant i\leqslant N_n} \mathrm{Inf}_i(f_n) \to 0$ *as* $n \to \infty$, *then*

$$\big|E\big[\psi(Q_d(N_n, f_n, \mathbf{X}))\big] - E\big[\psi(Q_d(N_n, f_n, \mathbf{G}))\big]\big| \to 0 \quad \text{as } n \to \infty, \quad (4.43)$$

for every continuous and bounded function $\psi : \mathbb{R} \to \mathbb{R}$.

3. *If $\beta := \sup_{i\geqslant 1} E[|X_i|^3] < \infty$ then, for all thrice differentiable $\varphi : \mathbb{R} \to \mathbb{R}$ such that $\|\varphi'''\|_\infty < \infty$ and for every fixed n,*

$$\big|E\big[\varphi(Q_d(N_n, f_n, \mathbf{X}))\big] - E\big[\varphi(Q_d(N_n, f_n, \mathbf{G}))\big]\big| \qquad (4.44)$$
$$\leqslant \|\varphi'''\|_\infty\,(30\beta)^d d!\,\sqrt{\max_{1\leqslant i\leqslant N_n} \mathrm{Inf}_i(f_n)}.$$

Proof. Point 1 is Theorem 2.2 in [20]. Point 2 is Proposition 1 in [40]. Point 3 is Theorem 3.18 (under Hypothesis H2) in [20]. Note that our polynomials Q_d relate to polynomials $d!Q$ in [20], hence the extra factor of $d!$ in 4.44. ■

Remark 4.2 We stress that Point 3 in Theorem 4.1 is a *non-asymptotic result*. In particular, the estimate (4.44) holds for every n independently of the asymptotic behavior of the quantities $\max_{1\leqslant i\leqslant N_n} \mathrm{Inf}_i(f_n)$.

In the sequel, we will also need the following technical lemma, which is a combination of several results stated and proved in [20].

Lemma 4.3 *Let $\mathbf{X} = \{X_i,\, i \geqslant 1\}$ be a collection of centered independent random variables with unit variance. Assume moreover that $\gamma := \sup_{i\geqslant 1} E[|X_i|^q] < \infty$ for some $q > 2$. Fix $N, d \geqslant 1$, and let $f : [N]^d \to \mathbb{R}$ be a symmetric function (here, observe that we do not require that f vanishes on diagonals). Define $Q_d(\mathbf{X}) = Q_d(N, f, \mathbf{X})$ by (1.1). Then*

$$E\big[|Q_d(\mathbf{X})|^q\big] \leqslant \gamma^d \big(2\sqrt{q-1}\big)^{qd} E\big[Q_d(\mathbf{X})^2\big]^{q/2}.$$

Proof. Combine Propositions 3.11, 3.16 and 3.12 in [20]. ■

As already evoked in the Introduction, one of the key elements in the proof of Theorem 4.1 given in [20] is the use of an elegant probabilistic technique, which is in turn inspired by the well-known Lindeberg's proof of the Central Limit Theorem (see [14]). Very lucid accounts of Lindeberg's ideas can be found in Trotter [46] and Feller [7, Section VIII]. As an illustration of the Lindeberg-type approach developed in [20], we will now state and prove a useful lemma, concerning moments of homogeneous sums. As suggested before, we believe that the employed technique may replace "diagram formulae" and similar combinatorial devices, that are customarily used in order to estimate moments of non-linear functionals of independent sequences (see e.g. [33, 43], as well as the estimates on moments of homogeneous sums given in de Jong's paper [6]).

Lemma 4.4 *Let $\mathbf{X} = \{X_i : i \geqslant 1\}$ and $\mathbf{Y} = \{Y_i : i \geqslant 1\}$ be two collections of centered independent random variables with unit variance. Fix some integers $N, d \geqslant 1$, and let $f : [N]^d \to \mathbb{R}$ be a symmetric function vanishing on diagonals. Define $Q_d(\mathbf{X}) = Q_d(N, f, \mathbf{X})$ and $Q_d(\mathbf{Y}) = Q_d(N, f, \mathbf{Y})$ according to (1.1).*

1. *Suppose $k \geqslant 2$ is such that: (a) X_i and Y_i belong to $L^k(\Omega)$ for all $i \geqslant 1$; (b) $E(X_i^l) = E(Y_i^l)$ for all $i \geqslant 1$ and $l \in \{2, \ldots, k\}$. Then $Q_d(\mathbf{X})$ and $Q_d(\mathbf{Y})$ belong to $L^k(\Omega)$, and $E(Q_d(\mathbf{X})^l) = E(Q_d(\mathbf{Y})^l)$ for all $l \in \{2, \ldots, k\}$.*

2. *Suppose $m > k \geqslant 2$ are such that: (a) $\alpha := \max\left\{\sup_{i\geqslant 1} E|X_i|^m, \sup_{i\geqslant 1} E|Y_i|^m\right\} < \infty$; (b) $E(X_i^l) = E(Y_i^l)$ for all $i \geqslant 1$ and $l \in \{2,\dots,k\}$. Assume moreover (for simplicity) that: (c) $E[Q_d(\mathbf{X})^2]^{1/2} \leqslant M$ for some finite constant $M \geqslant 1$. Then $Q_d(\mathbf{X})$ and $Q_d(\mathbf{Y})$ belong to $L^m(\Omega)$ and, for all $l \in \{k+1,\dots,m\}$,*

$$\begin{aligned}&\left|E(Q_d(\mathbf{X})^l) - E(Q_d(\mathbf{Y})^l)\right|\\ &\leqslant c_{d,l,m,\alpha} \times M^{l-k+1} \times \max_{1\leqslant i\leqslant N}\left\{\max\left[\mathrm{Inf}_i(f)^{\frac{k-1}{2}}; \mathrm{Inf}_i(f)^{\frac{l}{2}-1}\right]\right\},\end{aligned}$$

where $c_{d,l,m,\alpha} = \frac{2^{l+1}}{(d-1)!}\alpha^{\frac{dl}{m}}\left(2\sqrt{l-1}\right)^{(2d-1)l} d!^{l-1}$.

Proof. Point 1 could be verified by a direct (elementary) computation. However, we will obtain the same conclusion as the by-product of a more sophisticated construction (based on Lindeberg-type arguments) which will also lead to the proof of Point 2. We shall assume, without loss of generality, that the two sequences $\mathbf{X}$ and $\mathbf{Y}$ are stochastically independent. For $i = 0,\dots,N$, let $\mathbf{Z}^{(i)}$ denote the sequence $(Y_1,\dots,Y_i,X_{i+1},\dots,X_N)$. Fix a particular $i \in \{1,\dots,N\}$, and write

$$\begin{aligned}U_i &= \sum_{\substack{1\leqslant i_1,\dots,i_d\leqslant N\\ \forall k:\, i_k\neq i}} f(i_1,\dots,i_d) Z_{i_1}^{(i)}\dots Z_{i_d}^{(i)}\\ V_i &= \sum_{\substack{1\leqslant i_1,\dots,i_d\leqslant N\\ \exists k:\, i_k = i}} f(i_1,\dots,i_d) Z_{i_1}^{(i)}\dots \widehat{Z_i^{(i)}}\dots Z_{i_d}^{(i)},\end{aligned}$$

where $\widehat{Z_i^{(i)}}$ means that this particular term is dropped (observe that this notation bears no ambiguity: indeed, since f vanishes on diagonals, each string $i_1,...,i_d$ contributing to the definition of V_i contains the symbol i exactly once). Note that U_i and V_i are independent of the variables X_i and Y_i, and that $Q_d(\mathbf{Z}^{(i-1)}) = U_i + X_iV_i$ and $Q_d(\mathbf{Z}^{(i)}) = U_i + Y_iV_i$. By using the independence of X_i and Y_i from U_i and V_i (as well as the fact that $E(X_i^l) = E(Y_i^l)$ for all i and all $1 \leqslant l \leqslant k$) we infer from the binomial formula that, for $l \in \{2,\dots,k\}$,

$$\begin{aligned}E\big[(U_i + X_iV_i)^l\big] &= \sum_{j=0}^{l}\binom{l}{j} E(U_i^{l-j}V_i^j)E(X_i^j) \qquad (4.45)\\ &= \sum_{j=0}^{l}\binom{l}{j} E(U_i^{l-j}V_i^j)E(Y_i^j) = E\big[(U_i + Y_iV_i)^l\big].\end{aligned}$$

That is, $E\big[Q_d(\mathbf{Z}^{(i-1)})^l\big] = E\big[Q_d(\mathbf{Z}^{(i)})^l\big]$ for all $i \in \{1,\dots,N\}$ and $l \in \{2,\dots,k\}$. The desired conclusion of Point 1 follows by observing that $Q_d(\mathbf{Z}^{(0)}) = Q_d(\mathbf{X})$ and $Q_d(\mathbf{Z}^{(N)}) = Q_d(\mathbf{Y})$.

To prove Point 2, let $l \in \{k+1, \ldots, m\}$. Using (4.45) and then Hölder's inequality, we can write

$$\begin{aligned}
\big|E\big[Q_d(\mathbf{Z}^{(i-1)})^l\big] - E\big[Q_d(\mathbf{Z}^{(i)})^l\big]\big| &= \left|\sum_{j=k+1}^{l} \binom{l}{j} E(U_i^{l-j} V_i^j)\big(E(X_i^j) - E(Y_i^j)\big)\right| \\
&\leqslant \sum_{j=k+1}^{l} \binom{l}{j} (E|U_i|^l)^{1-j/l} (E|V_i|^l)^{j/l} \big(E|X_i|^j + E|Y_i|^j\big).
\end{aligned}$$

But, by Lemma 4.3 and since $E(U_i^2) \leqslant E(Q_d(\mathbf{X})^2) \leqslant M^2$, we have

$$E|U_i|^l \leqslant \alpha^{dl/m}\big(2\sqrt{l-1}\big)^{ld} E(U_i^2)^{l/2} \leqslant \alpha^{dl/m}\big(2\sqrt{l-1}\big)^{ld} M^l.$$

Similarly, since $E(V_i^2) = d!^2 \mathrm{Inf}_i(f)$ (see (1.8)), we have

$$E|V_i|^l \leqslant \alpha^{(d-1)l/m}\big(2\sqrt{l-1}\big)^{l(d-1)} E(V_i^2)^{l/2} \leqslant \alpha^{(d-1)l/m}\big(2\sqrt{l-1}\big)^{l(d-1)} d!^l \big(\mathrm{Inf}_i(f)\big)^{l/2}.$$

Hence, since $E|Y_i|^j + E|X_i|^j \leqslant 2\alpha^{j/m}$,

$$\begin{aligned}
&\big|E\big[Q_d(\mathbf{Z}^{(i-1)})^l\big] - E\big[Q_d(\mathbf{Z}^{(i)})^l\big]\big| \\
\leqslant\ & 2\sum_{j=k+1}^{l} \binom{l}{j} \Big(\alpha^{dl/m}\big(2\sqrt{l-1}\big)^{ld} M^l\Big)^{1-\frac{j}{l}} \Big(\alpha^{(d-1)/m}\big(2\sqrt{l-1}\big)^{d-1} d!\sqrt{\mathrm{Inf}_i(f)}\Big)^j \alpha^{j/m} \\
\leqslant\ & 2^{l+1}\alpha^{\frac{dl}{m}}\big(2\sqrt{l-1}\big)^{l(2d-1)} d!^l M^{l-k-1} \times \max\Big[\mathrm{Inf}_i(f)^{\frac{k+1}{2}}; \mathrm{Inf}_i(f)^{\frac{l}{2}}\Big].
\end{aligned}$$

Finally, summing for i over $1, \ldots, N$ and using that $\sum_{i=1}^N \mathrm{Inf}_i(f) = \frac{\|f\|_d^2}{(d-1)!} \leqslant \frac{M^2}{d!(d-1)!}$ yields

$$\begin{aligned}
&\big|E\big[Q_d(\mathbf{X})^l\big] - E\big[Q_d(\mathbf{Y})^l\big]\big| \\
\leqslant\ & 2^{l+1}\alpha^{\frac{dl}{m}}\big(2\sqrt{l-1}\big)^{l(2d-1)} d!^l M^{l-k-1} \\
&\qquad \times \max_{1\leqslant i\leqslant N}\Big\{\max\Big[\mathrm{Inf}_i(f)^{\frac{k-1}{2}}; \mathrm{Inf}_i(f)^{\frac{l}{2}-1}\Big]\Big\} \sum_{i=1}^N \mathrm{Inf}_i(f) \\
\leqslant\ & c_{d,l,m,\alpha} \times M^{l-k+1} \times \max_{1\leqslant i\leqslant N}\Big\{\max\Big[\mathrm{Inf}_i(f)^{\frac{k-1}{2}}; \mathrm{Inf}_i(f)^{\frac{l}{2}-1}\Big]\Big\}.
\end{aligned}$$

■

5 Normal approximation of homogeneous sums

5.1 Bounds

The following statement provides an explicit upper bound on the normal approximation of homogenous sums, when the test function has a bounded third derivative.

Theorem 5.1 *Let $\mathbf{X} = \{X_i,\, i \geqslant 1\}$ be a collection of centered independent random variables with unit variance. Assume moreover that $\sup_i E(X_i^4) < \infty$ and let $\alpha := \max\{3\,;\sup_i E(X_i^4)\}$. Fix $N, d \geqslant 1$, and let $f : [N]^d \to \mathbb{R}$ be symmetric and vanishing on diagonals. Define $Q_d(\mathbf{X}) = Q_d(N, f, \mathbf{X})$ according to (1.1) and assume that $E[Q_d(\mathbf{X})^2] = 1$. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a thrice differentiable function such that $\|\varphi'''\|_\infty \leqslant B$. Then, for $Z \sim \mathscr{N}(0,1)$, we have*

$$\big|E\big[\varphi(Q_d(\mathbf{X}))\big] - E\big[\varphi(Z)\big]\big| \leqslant B(30\beta)^d d! \sqrt{\max_{1\leqslant i\leqslant N} \mathrm{Inf}_i(f)} \tag{5.46}$$

$$+C_* \sqrt{\frac{d-1}{3d}} \left[\sqrt{\big|E\big[Q_d(\mathbf{X})^4\big] - 3\big|} + 4\sqrt{2} \times 144^{d-\frac{1}{2}}\, \alpha^{\frac{d}{2}} \sqrt{d} d! \left(\max_{1\leqslant i\leqslant N} \mathrm{Inf}_i(f)\right)^{1/4} \right]\Bigg),$$

with C_* defined as in Theorem 3.1.

Proof. Let $\mathbf{G} = (G_i)_{i\geqslant 1}$ be a standard centered i.i.d. Gaussian sequence. We have

$$\big|E\big[\varphi(Q_d(\mathbf{X}))\big] - E\big[\varphi(Z)\big]\big| \leqslant \delta_1 + \delta_2 \tag{5.47}$$

with $\delta_1 = \big|E\big[\varphi(Q_d(\mathbf{X}))\big] - E\big[\varphi(Q_d(\mathbf{G}))\big]\big|$ and $\delta_2 = \big|E\big[\varphi(Q_d(\mathbf{G}))\big] - E\big[\varphi(Z)\big]\big|$. By Theorem 4.1 point 3, we have $\delta_1 \leqslant B(30\beta)^d d! \sqrt{\max_{1\leqslant i\leqslant N} \mathrm{Inf}_i(f)}$. By Theorem 3.1 and since $E\big[Q_d(\mathbf{X})^2\big] = E\big[Q_d(\mathbf{G})^2\big] = 1$, we have

$$\delta_2 \leqslant C_* \sqrt{\frac{d-1}{3d} \big|E\big[Q_d(\mathbf{G})^4\big] - 3\big|}.$$

By Lemma 4.4 point 2 (with $M = 1$, $k = 2$ and $l = m = 4$) and since $\mathrm{Inf}_i(f) \leqslant 1$ for all i, we have

$$\big|E\big[Q_d(\mathbf{X})^4\big] - E\big[Q_d(\mathbf{G})^4\big]\big| \leqslant 32 \times 144^{2d-1}\, \alpha^d d d!^2 \sqrt{\max_{1\leqslant i\leqslant N} \mathrm{Inf}_i(f)} \tag{5.48}$$

so that

$$\delta_2 \leqslant C_* \sqrt{\frac{d-1}{3d}} \left[\sqrt{\big|E\big[Q_d(\mathbf{X})^4\big] - 3\big|} + 4\sqrt{2} \times 144^{d-\frac{1}{2}}\, \alpha^{\frac{d}{2}} \sqrt{d} d! \left(\max_{1\leqslant i\leqslant N} \mathrm{Inf}_i(f)\right)^{1/4} \right].$$

Finally, the desired conclusion follows by putting the bounds for δ_1 and δ_2 in (5.47). ∎

Remark 5.2 As a corollary of Theorem 5.1, we immediately recover de Jong's Theorem 1.10, under the additional hypothesis that $\sup_i E(X_i^4) < \infty$.

5.2 Converse statements: universality of Wiener chaos

Here, we prove a slightly stronger version of Theorem 1.12 stated in Section 1.5. More precisely, we have the following result, where an additional condition on contractions (see assumption (3) in Theorem 5.3 just below and Definition 1.6) have been added with respect to Theorem 1.12, making the criterion more easily applicable in practice.

Theorem 5.3 *Let $\mathbf{G} = \{G_i : i \geqslant 1\}$ be a standard centered i.i.d. Gaussian sequence, and fix an even integer $d \geqslant 2$. Let $\{N_n, f_n : n \geqslant 1\}$ be a sequence such that $\{N_n\}$ is a sequence of integers going to infinity, and each $f_n : [N_n]^d \to \mathbb{R}$ is symmetric and vanishes on diagonals. Define $Q_d(n, \mathbf{G}) = Q_d(N_n, f_n, \mathbf{G})$, $n \geqslant 1$, according to (1.1), and assume that $E[Q_d(n, \mathbf{G})^2] \to 1$ as $n \to \infty$. Then, the following are equivalent as $n \to \infty$.*

(1) *The sequence $Q_d(n, \mathbf{G})$ converges in law to $Z \sim \mathscr{N}(0, 1)$.*

(2) $E[Q_d(n, \mathbf{G})^4] \to 3$.

(3) *For all $r = 1, \ldots, d-1$, $\|f_n \star_r f_n\|_{2d-2r} \to 0$.*

(4) *For every sequence $\mathbf{X} = \{X_i : i \geqslant 1\}$ of independent centered random variables with unit variance and such that $\sup_i E|X_i|^{2+\epsilon} < \infty$ for some $\epsilon > 0$, the sequence $Q_d(n, \mathbf{X})$ converges in law to $Z \sim \mathscr{N}(0, 1)$, in the Kolmogorov distance.*

(5) *For every sequence $\mathbf{X} = \{X_i : i \geqslant 1\}$ of independent and identically distributed centered random variables with unit variance, the sequence $Q_d(n, \mathbf{X})$ converges in law to $Z \sim \mathscr{N}(0, 1)$ (not necessarily in the Kolmogorov distance).*

Proof. The equivalences (1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3) are a mere reformulation of Theorem 3.3, deduced by taking into account the equality (3.36). On the other hand, it is trivial that each one of conditions (4) and (5) implies (1). So, it remains to prove the implication $(1), (2), (3) \Rightarrow (4), (5)$. Fix $z \in \mathbb{R}$. We have

$$\big|P\big[Q_d(n, \mathbf{X}) \leqslant z\big] - P\big[Z \leqslant z\big]\big| \leqslant \delta_n^{(a)}(z) + \delta_n^{(b)}(z)$$

with

$$\begin{aligned}
\delta_n^{(a)}(z) &= \big|P\big[Q_d(n, \mathbf{X}) \leqslant z\big] - P\big[Q_d(n, \mathbf{G}) \leqslant z\big]\big| \\
\delta_n^{(b)}(z) &= \big|P\big[Q_d(n, \mathbf{G}) \leqslant z\big] - P\big[Z \leqslant z\big]\big|.
\end{aligned}$$

By assumption (2) and (3.30)-(3.31), we have $\sup_{z\in\mathbb{R}} \delta_n^{(b)}(z) \to 0$. By combining assumption (3) (for $r = d-1$) with (1.14)-(1.15), we get that $\max_{1\leqslant i\leqslant N_n} \mathrm{Inf}_i(f_n) \to 0$ as $n \to \infty$. Hence, Theorem 4.1 (Point 1) implies that $\sup_{z\in\mathbb{R}} \delta_n^{(a)}(z) \to 0$, and the proof of the implication $(1), (2), (3) \Rightarrow (4)$ is complete. To prove that $(1) \Rightarrow (5)$, one uses the same line of reasoning, the only difference being that we need to use Point 2 of Theorem 4.1 instead of Point 1. ■

5.3 Bounds on the Wasserstein distance

The following statement shows that our techniques allow to directly control the Wasserstein distance between the law of a homogenous sum and the law of a standard Gaussian random variable.

Proposition 5.4 *As in Theorem 5.1, let* $\mathbf{X} = \{X_i,\, i \geqslant 1\}$ *be a collection of centered independent random variables with unit variance. Assume moreover that* $\sup_i E(X_i^4) < \infty$ *and note* $\alpha := \max\{3\,;\, \sup_i E(X_i^4)\}$*. Fix* $N, d \geqslant 1$*, and let* $f : [N]^d \to \mathbb{R}$ *be symmetric and vanishing on diagonals. Define* $Q_d(\mathbf{X}) = Q_d(N, f, \mathbf{X})$ *according to (1.1) and assume that* $E[Q_d(\mathbf{X})^2] = 1$*. Put*

$$B_1 = 2(30\beta)^d d! \sqrt{\max_{1\leqslant i\leqslant N} \mathrm{Inf}_i(f)}$$

and

$$B_2 = 3 \times \left(3 + (2\sqrt{2})^d\right)\sqrt{\frac{d-1}{3d}} \times$$
$$\times \left[\sqrt{\left|E\left[Q_d(\mathbf{X})^4\right] - 3\right|} + 4\sqrt{2} \times 144^{d-\frac{1}{2}}\, \alpha^{\frac{d}{2}} \sqrt{d}d! \left(\max_{1\leqslant i\leqslant N} \mathrm{Inf}_i(f)\right)^{\frac{1}{4}} \right].$$

Then, for $Z \sim \mathscr{N}(0,1)$*, we have* $d_W(Q_d(\mathbf{X}), Z) \leqslant 4(B_1 + B_2)^{\frac{1}{3}}$*, provided* $B_1 + B_2 \leqslant \frac{3}{4\sqrt{2}}$*.*

Proof. Let $h \in \mathrm{Lip}(1)$ be a Lipschitz function with Lipschitz constant 1. By Rademacher's theorem, h is Lebesgue-almost everywhere differentiable. So, if we denote by h' its derivative, then $\|h'\|_\infty \leqslant 1$. For $t > 0$, define

$$h_t(x) = \int_{-\infty}^{\infty} h(\sqrt{t}y + \sqrt{1-t}x)\phi(y)dy,$$

where ϕ denotes the standard normal density. The triangle inequality gives

$$\begin{aligned} &\left|E[h(Q_d(\mathbf{X}))] - E[h(Z)]\right| \\ \leqslant\; &\left|E[h_t(Q_d(\mathbf{X}))] - E[h_t(Z)]\right| + \left|E[h(Q_d(\mathbf{X}))] - E[h_t(Q_d(\mathbf{X}))]\right| + \left|E[h(Z)] - E[h_t(Z)]\right|. \end{aligned}$$

Now, let us differentiate and integrate by parts, using that $\phi'(x) = -x\phi(x)$, to get

$$h_t''(x) = \frac{1-t}{\sqrt{t}} \int_{-\infty}^{\infty} yh'(\sqrt{t}y + \sqrt{1-t}x)\phi\,(y)\,dy.$$

Hence, for $0 < t < 1$, we may bound

$$\|h_t''\|_\infty \leqslant \frac{1-t}{\sqrt{t}}\|h'\|_\infty \int_{-\infty}^{\infty} |y|\phi(y)dy \leqslant \frac{1}{\sqrt{t}}. \tag{5.49}$$

For $0 < t \leqslant \frac{1}{2}$ (so that $\sqrt{t} \leqslant \sqrt{1-t}$), we have

$$
\begin{aligned}
& |E[h(Q_d(\mathbf{X}))] - E[h_t(Q_d(\mathbf{X}))]| \\
\leqslant\ & \left| E\left[\int_{-\infty}^{\infty} \left\{ h(\sqrt{t}y + \sqrt{1-t}Q_d(\mathbf{X})) - h(\sqrt{1-t}Q_d(\mathbf{X})) \right\} \phi(y)dy \right] \right| \\
& + E\left[\left| h(\sqrt{1-t}Q_d(\mathbf{X})) - h(Q_d(\mathbf{X})) \right| \right] \\
\leqslant\ & \|h'\|_\infty \sqrt{t} \int_{-\infty}^{\infty} |y|\phi(y)dy + \|h'\|_\infty \frac{t}{2\sqrt{1-t}} E\left[|Q_d(\mathbf{X})| \right] \leqslant \frac{3}{2}\sqrt{t}.
\end{aligned}
$$

Similarly, $|E[h(Z)] - E[h_t(Z)]| \leqslant \frac{3}{2}\sqrt{t}$.

We now apply Theorem 5.1. To bound C_*, we use that $|h_t'(0)| \leqslant 1$ and that $|h_t''(0)| \leqslant t^{-\frac{1}{2}}$; also $\|h_t'''\|_\infty \leqslant 2/t$ (as it can be shown by using the same arguments as in the proof of (5.49) above). Hence, as $2 \leqslant \frac{1}{t}$ and $\sqrt{2} \leqslant t^{-\frac{1}{2}}$, we have

$$
\begin{aligned}
C_* = \left(3 + (2\sqrt{2})^d\right) \times K_* \ &\leqslant\ \left(3 + (2\sqrt{2})^d\right) \times \max\left\{ 2t^{-\frac{1}{2}} + \frac{4\sqrt{2}}{3\sqrt{\pi}} t^{-1}; 2 + \frac{2}{3} t^{-1} \right\} \\
&\leqslant\ \left(3 + (2\sqrt{2})^d\right) \times \frac{3}{t}.
\end{aligned}
$$

Due to $\|h_t'''\|_\infty \leqslant 2/t$, Theorem 5.1 gives the bound

$$
\left| E\left[h_t(Q_d(\mathbf{X}))\right] - E\left[h_t(Z)\right] \right| \ \leqslant\ 3\sqrt{t} + (B_1 + B_2)\frac{1}{t}.
$$

Minimising $3\sqrt{t} + (B_1 + B_2)\frac{1}{t}$ in t gives that $t = \left(\frac{2}{3}(B_1 + B_2)\right)^{\frac{2}{3}}$. Plugging in the values and bounding the constant part ends the proof. ■

6 Chi-square approximation of homogeneous sums

6.1 Bounds

The next result provides upper bounds on the chi-square approximation of homogeneous sums.

Theorem 6.1 *Let $\mathbf{X} = \{X_i, i \geqslant 1\}$ be a collection of centered independent random variables with unit variance. Assume moreover that $\sup_i E(X_i^4) < \infty$ and note $\alpha := \max\{3; \sup_i E(X_i^4)\}$. Fix an* even *integer $d \geqslant 2$ and, for $N \geqslant 1$, let $f : [N]^d \to \mathbb{R}$ be symmetric and vanishing on diagonals. Define $Q_d(\mathbf{X}) = Q_d(N, f, \mathbf{X})$ according to (1.1) and assume that $E[Q_d(\mathbf{X})^2] = 2\nu$ for some integer $\nu \geqslant 1$. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a thrice differentiable function such that $\|\varphi\|_\infty \leqslant 1$, $\|\varphi'\|_\infty \leqslant 1$ and $\|\varphi'''\|_\infty \leqslant B$. Then, for*

$Z_\nu \sim \chi^2(\nu)$, *we have*

$$\begin{aligned}
&\big|E\big[\varphi(Q_d(\mathbf{X}))\big] - E\big[\varphi(Z_\nu)\big]\big| \\
\leqslant\ & B(30\beta)^d d! \sqrt{\max_{1\leqslant i\leqslant N} \mathrm{Inf}_i(f)} + \max\left\{\sqrt{\frac{2\pi}{\nu}}, \frac{1}{\nu} + \frac{2}{\nu^2}\right\} \\
&\times \left(\sqrt{\frac{d-1}{3d}}\left[\sqrt{\big|E\big[Q_d(\mathbf{X})^4\big] - 12E\big[Q_d(\mathbf{X})^3\big] - 12\nu^2 + 48\nu\big|}\right.\right. \\
&\left.\left. +4\sqrt{d}d!\left(\sqrt{2}\times 144^{d-\frac{1}{2}}\alpha^{\frac{d}{2}} + \sqrt{\nu}(2\sqrt{2})^{\frac{3(2d-1)}{2}}\alpha^{\frac{3d}{2}}\right)\left(\max_{1\leqslant i\leqslant N}\mathrm{Inf}_i(f)\right)^{1/4}\right]\right).
\end{aligned}$$

Proof. We proceed as in Theorem 5.1. Let $\mathbf{G} = (G_i)_{i\geqslant 1}$ denote a standard centered i.i.d. Gaussian sequence. We have $\big|E\big[\varphi(Q_d(\mathbf{X}))\big] - E\big[\varphi(Z_\nu)\big]\big| \leqslant \delta_1 + \delta_2$ with

$$\begin{aligned}
\delta_1 &= \big|E\big[\varphi(Q_d(\mathbf{X}))\big] - E\big[\varphi(Q_d(\mathbf{G}))\big]\big| \\
\delta_2 &= \big|E\big[\varphi(Q_d(\mathbf{G}))\big] - E\big[\varphi(Z_\nu)\big]\big|.
\end{aligned}$$

By Theorem 4.1 (Point 3), we have $\delta_1 \leqslant B(30\beta)^d d!\sqrt{\max_{1\leqslant i\leqslant N}\mathrm{Inf}_i(f)}$. By Theorem 3.6, we have, with $C_\# = \max\left\{\sqrt{\frac{2\pi}{\nu}}, \frac{1}{\nu}+\frac{2}{\nu^2}\right\}$:

$$\delta_2 \leqslant d_{BW}(Q_d(\mathbf{G}), Z_\nu) \leqslant C_\# \sqrt{\frac{d-1}{3d}\big|E\big[Q_d(\mathbf{G})^4\big] - 12E\big[Q_d(\mathbf{G})^3\big] - 12\nu^2 + 48\nu\big|}.$$

Additionally to (5.48), we have, by Lemma 4.4 (Point 2, with $M = \sqrt{2\nu}$, $k=2$, $l=3$ and $m=4$):

$$\big|E\big[Q_d(\mathbf{X})^3\big] - E\big[Q_d(\mathbf{G})^3\big]\big| \leqslant 16\nu(2\sqrt{2})^{3(2d-1)}\alpha^{\frac{3d}{4}} d d! \sqrt{\max_{1\leqslant i\leqslant N}\mathrm{Inf}_i(f)}.$$

Hence

$$\begin{aligned}
\delta_2 \leqslant\ & C_\# \sqrt{\frac{d-1}{3d}}\left[\sqrt{\big|E\big[Q_d(\mathbf{X})^4\big] - 12E\big[Q_d(\mathbf{X})^3\big] - 12\nu^2 + 48\nu\big|}\right. \\
&\left. +4\sqrt{d}d!\left(\sqrt{2}\times 144^{d-\frac{1}{2}}\alpha^{\frac{d}{2}} + \sqrt{\nu}\,(2\sqrt{2})^{\frac{3(2d-1)}{2}}\alpha^{\frac{3d}{2}}\right)\left(\max_{1\leqslant i\leqslant N}\mathrm{Inf}_i(f)\right)^{1/4}\right].
\end{aligned}$$

The desired conclusion follows. ■

As an immediate corollary of Theorem 6.1, we deduce the following new criterion for the asymptotic non-normality of homogenous sums – compare with de Jong's Theorem 1.10.

Corollary 6.2 *Let* $\mathbf{X} = \{X_i : i \geqslant 1\}$ *be a sequence of independent centered random variables with unit variance such that* $\sup_i E(X_i^4) < \infty$. *Fix an* even *integer* $d \geqslant 2$, *and let* $\{N_n, f_n : n \geqslant 1\}$ *be a sequence such that* $\{N_n : n \geqslant 1\}$ *is a sequence of integers going to infinity, and each* $f_n : [N_n]^d \to \mathbb{R}$ *is symmetric and vanishes on diagonals. Define* $Q_d(n, \mathbf{X}) = Q_d(N_n, f_n, \mathbf{X})$ *according to (1.1). If, as* $n \to \infty$,

(i) $E(Q_d(n,\mathbf{X})^2) \to 2\nu$,

(ii) $E[Q_d(n,\mathbf{X})^4] - 12E[Q_d(N_n, f_n, \mathbf{X})^3] \to 12\nu^2 - 48\nu$, *and*

(iii) $\max_{1\leqslant i\leqslant N_n} \mathrm{Inf}_i(f_n) \to 0$,

then $Q_d(n,\mathbf{X})$ *converges in law to* $Z_\nu \sim \chi^2(\nu)$.

6.2 A converse statement

The following statement contains a universal chi-square limit theorem result: it is a general version of Theorem 1.14.

Theorem 6.3 *Let* $\mathbf{G} = \{G_i : i \geqslant 1\}$ *be a standard centered i.i.d. Gaussian sequence, and fix an* even *integer* $d \geqslant 2$. *Let* $\{N_n, f_n : n \geqslant 1\}$ *be a sequence such that* $\{N_n\}$ *is a sequence of integers going to infinity, and each* $f_n : [N_n]^d \to \mathbb{R}$ *is symmetric and vanishes on diagonals. Define* $Q_d(n,\mathbf{G}) = Q_d(N_n, f_n, \mathbf{G})$, $n \geqslant 1$, *according to (1.1), and assume that* $E[Q_d(n,\mathbf{G})^2] \to 2\nu$ *as* $n \to \infty$. *Then, the following are equivalent as* $n \to \infty$.

(1) *The sequence* $Q_d(n,\mathbf{G})$ *converges in law to* $Z_\nu \sim \chi^2(\nu)$.

(2) $E[Q_d(n,\mathbf{G})^4] - 12E[Q_d(n,\mathbf{G})^3] \to 12\nu^2 - 48\nu$.

(3) $\|f_n \widetilde{\star}_{d/2} f_n - c_d \times f_n\|_d \to 0$ *(for* c_d *defined in (1.17)) and, for every* $r = 1, ..., d-1$ *such that* $r \neq d/2$, $\|f_n \star_r f_n\|_{2d-2r} \to 0$.

(4) *For every sequence* $\mathbf{X} = \{X_i : i \geqslant 1\}$ *of independent centered random variables with unit variance and such that* $\sup_i E|X_i|^{2+\epsilon} < \infty$ *for some* $\epsilon > 0$, *the sequence* $Q_d(n,\mathbf{X})$ *converges in law to* $Z_\nu \sim \chi^2(\nu)$.

(5) *For every sequence* $\mathbf{X} = \{X_i : i \geqslant 1\}$ *of independent and identically distributed centered random variables with unit variance, the sequence* $Q_d(n,\mathbf{X})$ *converges in law to* Z_ν.

Proof. The proof follows exactly the same lines of reasoning as in Theorem 5.3. Details are left to the reader. Let us just mention that the only differences consist in the use Theorem 3.8 instead of Theorem 3.3, and the use of (3.39)–(3.40) instead of (3.30)–(3.31). ■

7 Multivariate Extensions

7.1 Bounds

We recall here the standard multi-index notation. A multi-index is a vector $\alpha \in \{0, 1, \ldots\}^m$. We write

$$|\alpha| = \sum_{j=1}^m \alpha_j, \quad \alpha! = \prod_{j=1}^m \alpha_j!, \quad \partial_j = \frac{\partial}{\partial x_j}, \quad \partial^\alpha = \partial_1^{\alpha_1} \ldots \partial_d^{\alpha_d}, \quad x^\alpha = \prod_{j=1}^m x_j^{\alpha_j}.$$

Note that by convention, $0^0 = 1$. Also note that $|x^\alpha| = y^\alpha$, where $y_j = |x_j|$ for all j. Finally, for $\varphi : \mathbb{R}^m \to \mathbb{R}$ regular enough and $k \geqslant 1$, we put

$$\|\varphi^{(k)}\|_\infty = \max_{|\alpha|=k} \frac{1}{\alpha!} \sup_{z\in\mathbb{R}^m} |\partial^\alpha \varphi(z)|.$$

First we derive a multivariate version of (4.44); see also [19, Theorem 4.1] for a similar statement.

Theorem 7.1 *Let $\mathbf{X} = \{X_i,\ i \geqslant 1\}$ be a collection of centered independent random variables with unit variance and such that $\beta := \sup_{i\geqslant 1} E[|X_i|^3] < \infty$. Let $\mathbf{G} = \{G_i : i \geqslant 1\}$ be a standard centered i.i.d. Gaussian sequence. Fix integers $m \geqslant 1$ and $d_m \geqslant \ldots \geqslant d_1 \geqslant 1$ as well as $N_1, \ldots, N_m \geqslant 1$. For every $j = 1, ..., m$, let $f_j : [N_j]^{d_j} \to \mathbb{R}$ be a symmetric function vanishing on diagonals. Define $Q^j(\mathbf{G}) = Q_{d_j}(N_j, f_j, \mathbf{G})$ and $Q^j(\mathbf{X}) = Q_{d_j}(N_j, f_j, \mathbf{X})$ according to (1.1), and assume that $E[Q^j(\mathbf{G})^2] = E[Q^j(\mathbf{X})^2] = 1$ for all $j = 1, \ldots, m$. Assume that there exists a $C > 0$ is such that $\sum_{i=1}^{\max_j N_j} \max_{1\leqslant j\leqslant m} \mathrm{Inf}_i(f_j) \leqslant C$. Then, for all thrice differentiable $\varphi : \mathbb{R}^m \to \mathbb{R}$ with $\|\varphi'''\|_\infty < \infty$, we have*

$$\begin{aligned} &\big|E\big[\varphi(Q^1(\mathbf{X}), \ldots, Q^m(\mathbf{X}))\big] - E\big[\varphi(Q^1(\mathbf{G}), \ldots, Q^m(\mathbf{G}))\big]\big| \\ \leqslant\ & C\|\varphi'''\|_\infty \left(\beta + \sqrt{\frac{8}{\pi}}\right) \left[\sum_{j=1}^m (16\sqrt{2}\beta)^{\frac{d_j-1}{3}} d_j!\right]^3 \sqrt{\max_{1\leqslant j\leqslant m} \max_{1\leqslant i\leqslant \max_j N_j} \mathrm{Inf}_i(f_j)}. \end{aligned}$$

In the one-dimensional case (i.e. $m = 1$), $\sum_{i=1}^{\max_j N_j} \max_{1\leqslant j\leqslant m} \mathrm{Inf}_i(f_j) = [d!(d-1)!]^{-1}$, so we can choose $C = [d!(d-1)!]^{-1}$. In this case, when β is large the bound from Theorem 7.1 essentially differs from (4.44) by a constant times a factor d.

Proof. Abbreviate

$$\mathbf{Q}(\mathbf{X}) = \big(Q^1(\mathbf{X}), \ldots, Q^m(\mathbf{X})\big),$$

and define analogously $\mathbf{Q}(\mathbf{G})$. We proceed as for Lemma 4.4, with similar notation. For $i = 0, \ldots, \max_j N_j$, let $\mathbf{Z}^{(i)}$ denote the sequence $\big(G_1, \ldots, G_i, X_{i+1}, \ldots, X_{\max_j N_j}\big)$. Using the triangle inequality,

$$\Big|E\big[\varphi(\mathbf{Q}(\mathbf{X}))\big] - E\big[\varphi(\mathbf{Q}(\mathbf{G}))\big]\Big| \leqslant \sum_{i=1}^{\max_j N_j} \Big|E\big[\varphi(\mathbf{Q}(\mathbf{Z}^{(i-1)}))\big] - E\big[\varphi(\mathbf{Q}(\mathbf{Z}^{(i)}))\big]\Big|.$$

Fix a particular $i \in \{1, \ldots, \max_j N_j\}$, and write, for $j = 1, \ldots, m$,

$$\begin{aligned} U_i^{(j)} &= \sum_{\substack{1\leqslant i_1,\ldots,i_d\leqslant N_j \\ \forall k:\, i_k\neq i}} f_j(i_1, \ldots, i_d) Z_{i_1}^{(i)} \ldots Z_{i_d}^{(i)} \\ V_i^{(j)} &= \sum_{\substack{1\leqslant i_1,\ldots,i_d\leqslant N_j \\ \exists k:\, i_k=i}} f_j(i_1, \ldots, i_d) Z_{i_1}^{(i)} \ldots \widehat{Z_i^{(i)}} \ldots Z_{i_d}^{(i)}, \end{aligned}$$

where $\widehat{Z_i^{(i)}}$ means that this particular term is dropped. We write $\mathbf{U}_i = (U_i^{(1)}, \ldots, U_i^{(m)})$ and $\mathbf{V}_i = (V_i^{(1)}, \ldots, V_i^{(m)})$. Note that $\mathbf{U}_i$ and $\mathbf{V}_i$ are independent of the variables X_i and G_i, and that $\mathbf{Q}(\mathbf{Z}^{(i-1)}) = \mathbf{U}_i + X_i\mathbf{V}_i$ and $\mathbf{Q}(\mathbf{Z}^{(i)}) = \mathbf{U}_i + G_i\mathbf{V}_i$. By Taylor's theorem, using the independence of X_i from $\mathbf{U}_i$ and $\mathbf{V}_i$ as well as $E(X_i) = 0$ and $E(X_i^2) = 1$, we have

$$\begin{aligned}&\left| E\big[\varphi(\mathbf{U}_i + X_i\mathbf{V}_i)\big] - \sum_{|\alpha|\leqslant 2} \frac{1}{\alpha!} E\big(\partial^\alpha \varphi(\mathbf{U}_i)\mathbf{V}_i^\alpha\big) E(X_i^{|\alpha|}) \right| \\ \leqslant \; & \sum_{|\alpha|=3} \frac{1}{\alpha!} \sup_{z\in\mathbb{R}^m} |\partial^\alpha \varphi(z)|\, E(|X_i|^3) E(|\mathbf{V}_i^\alpha|) \leqslant \|\varphi'''\|_\infty E(|X_i|^3) \sum_{|\alpha|=3} E(|\mathbf{V}_i^\alpha|).\end{aligned}$$

Similarly,

$$\begin{aligned}&\left| E\big[\varphi(\mathbf{U}_i + G_i\mathbf{V}_i)\big] - \sum_{|\alpha|\leqslant 2} \frac{1}{\alpha!} E\big(\partial^\alpha \varphi(\mathbf{U}_i)\mathbf{V}_i^\alpha\big) E(G_i^{|\alpha|}) \right| \\ \leqslant \; & \sqrt{\frac{8}{\pi}} \sum_{|\alpha|=3} \frac{1}{\alpha!} \sup_{z\in\mathbb{R}^m} |\partial^\alpha \varphi(z)|\, E(|\mathbf{V}_i^\alpha|) \leqslant \|\varphi'''\|_\infty \sqrt{\frac{8}{\pi}} \sum_{|\alpha|=3} E(|\mathbf{V}_i^\alpha|).\end{aligned}$$

Due to the independence and the matching moments up to second order, we obtain that

$$\begin{aligned}&\big|E\big[\varphi(\mathbf{Q}(n, \mathbf{Z}^{(i-1)}))\big] - E\big[\varphi(\mathbf{Q}(n, \mathbf{Z}^{(i)}))\big]\big| = \big|E\big[\varphi(\mathbf{U}_i + X_i\mathbf{V}_i)\big] - E\big[\varphi(\mathbf{U}_i + G_i\mathbf{V}_i)\big]\big| \\ \leqslant \; & \left(\beta + \sqrt{\frac{8}{\pi}}\right) \|\varphi'''\|_\infty \sum_{|\alpha|=3} E(|\mathbf{V}_i^\alpha|).\end{aligned}$$

Abbreviate $\tau_i = \max_{1\leqslant j\leqslant m} \mathrm{Inf}_i(f_j)$. Next we use that, for $j = 1, \ldots, m$, by Lemma 4.3 (with $q = 3$), we have

$$E[|V_i^{(j)}|^3] \leqslant (16\sqrt{2}\beta)^{d_j - 1}\, E[(V_i^{(j)})^2]^{\frac{3}{2}} = (16\sqrt{2}\beta)^{d_j-1} d_j!^3 \tau_i^{3/2}.$$

Thus

$$\begin{aligned}\sum_{|\alpha|=3} E\,|(\mathbf{V}_i)^\alpha| &= \sum_{j,k,l=1}^m E(|V_i^{(j)} V_i^{(k)} V_i^{(l)}|) \leqslant \sum_{j,k,l=1}^m E(|V_i^{(j)}|^3)^{\frac{1}{3}} E(|V_i^{(k)}|^3)^{\frac{1}{3}} E(|V_i^{(l)}|^3)^{\frac{1}{3}} \\ &= \left(\sum_{j=1}^m E(|V_i^{(j)}|^3)^{\frac{1}{3}}\right)^3 \leqslant \left[\sum_{j=1}^m (16\sqrt{2}\beta)^{\frac{d_j-1}{3}} d_j!\right]^3 \tau_i^{3/2}.\end{aligned}$$

Collecting the bounds, summing over i, and using that $\sum_{i=1}^{\max_j N_j} \tau_i \leqslant C$ gives the desired result.

■

The next theorem associates bounds to the normal approximation of the vector $(Q^1(\mathbf{X}), \ldots, Q^m(\mathbf{X}))$.

Theorem 7.2 *Let $\mathbf{X} = \{X_i : i \geqslant 1\}$ be a collection of centered independent random variables with unit variance. Assume moreover that $\beta := \sup_i E\big[|X_i|^3\big] < \infty$. Fix integers $m \geqslant 1$, $d_m \geqslant \ldots \geqslant d_1 \geqslant 2$ and $N_1, \ldots, N_m \geqslant 1$. For every $j = 1, ..., m$, let $f_j : [N_j]^{d_j} \to \mathbb{R}$ be a symmetric function vanishing on diagonals. Define $Q^j(\mathbf{X}) = Q_{d_j}(N_j, f_j, \mathbf{X})$ according to (1.1), and assume that $E[Q^j(\mathbf{X})^2] = 1$ for all $j = 1, \ldots, m$. Let V be the $m \times m$ symmetric matrix given by $V(i,j) = E\big[Q^i(\mathbf{X})Q^j(\mathbf{X})\big]$. Let C be as in Theorem 7.1. Let $\varphi : \mathbb{R}^3 \to \mathbb{R}$ be a thrice differentiable function such that $\|\varphi''\|_\infty < \infty$ and $\|\varphi'''\|_\infty < \infty$. Then, for $Z_V = (Z_V^1, \ldots, Z_V^m) \sim \mathscr{N}_m(0, V)$ (centered Gaussian vector with covariance matrix V), we have*

$$\big|E\big[\varphi(Q^1(\mathbf{X}), \ldots, Q^m(\mathbf{X}))\big] - E\big[\varphi(Z_V)\big]\big| \leqslant \|\varphi''\|_\infty \left(\sum_{i=1}^m \Delta_{ii} + 2\sum_{1\leqslant i<j\leqslant m} \Delta_{ij}\right)$$
$$+C\|\varphi'''\|_\infty \left(\beta + \sqrt{\frac{8}{\pi}}\right)\left[\sum_{j=1}^m (16\sqrt{2}\beta)^{\frac{d_j-1}{3}} d_j!\right]^3 \sqrt{\max_{1\leqslant j\leqslant m}\max_{1\leqslant i\leqslant N_j} \mathrm{Inf}_i(f_j)},$$

for Δ_{ij} given by

$$\frac{d_j}{\sqrt{2}}\sum_{r=1}^{d_i-1}(r-1)!\binom{d_i-1}{r-1}\binom{d_j-1}{r-1}\sqrt{(d_i+d_j-2r)!}\,\big(\|f_i \star_{d_i-r} f_i\|_{\mathfrak{H}^{\otimes(2r)}} + \|f_j \star_{d_j-r} f_j\|_{\mathfrak{H}^{\otimes(2r)}}\big)$$
$$+\mathbf{1}_{\{d_i<d_j\}}\sqrt{d_j!\binom{d_j}{d_i}\|f_j \star_{d_j-d_i} f_j\|_{\mathfrak{H}^{\otimes 2d_i}}}. \tag{7.50}$$

Proof. The proof is divided into several steps.

Step 1: Reduction of the problem. Let $\mathbf{G} = (G_i)_{i\geqslant 1}$ be a standard centered i.i.d. Gaussian sequence. We have

$$\big|E\big[\varphi(Q^1(\mathbf{X}), \ldots, Q^m(\mathbf{X}))\big] - E\big[\varphi(Z_V)\big]\big| \leqslant \delta_1 + \delta_2 \tag{7.51}$$

with

$$\begin{aligned}\delta_1 &= \big|E\big[\varphi(Q^1(\mathbf{X}), \ldots, Q^m(\mathbf{X}))\big] - E\big[\varphi(Q^1(\mathbf{G}), \ldots, Q^m(\mathbf{G}))\big]\big|\\ \delta_2 &= \big|E\big[\varphi(Q^1(\mathbf{G}), \ldots, Q^m(\mathbf{G}))\big] - E\big[\varphi(Z_V)\big]\big|.\end{aligned}$$

Step 2: Bounding δ_1. By Theorem 7.1, we have

$$\delta_1 \leqslant C\|\varphi'''\|_\infty \left(\beta + \sqrt{\frac{8}{\pi}}\right)\left[\sum_{j=1}^m (16\sqrt{2}\beta)^{\frac{d_j-1}{3}} d_j!\right]^3 \sqrt{\max_{1\leqslant j\leqslant m}\max_{1\leqslant i\leqslant N_j} \mathrm{Inf}_i(f_j)}.$$

Step 3: Bounding δ_2. We will not use the result proved in [24], since here we do not assume that the matrix V is positive definite. Instead, we will rather use an interpolation technique *à la* Talagrand [44]. Without loss of generality, we assume in this step that Z_V is independent of $\mathbf{G}$. Also, for any $j = 1, \dots, m$, observe that $Q^j(\mathbf{G}) = I_{d_j}(h_j)$ where

$$h_j = d_j! \sum_{\{i_1,\dots,i_{d_j}\}\subset[N_j]^{d_j}} f_j(i_1,\dots,i_{d_j})\, e_{i_1}\otimes\ldots\otimes e_{i_{d_j}} \in \mathfrak{H}^{\odot d},$$

see also (2.26). For $t\in[0,1]$, set

$$\Psi(t) = E\big[\varphi\big(\sqrt{1-t}(I_{d_1}(h_1),\dots,I_{d_m}(h_m)) + \sqrt{t}Z_V\big)\big],$$

so that

$$\delta_2 = \big|\Psi(1)-\Psi(0)\big| \leqslant \sup_{t\in(0,1)}\big|\Psi'(t)\big|.$$

We easily see that Ψ is differentiable on $(0,1)$, and that $\Psi'(t)$ equals

$$\sum_{i=1}^m E\left[\frac{\partial\varphi}{\partial x_i}\big(\sqrt{1-t}(I_{d_1}(h_1),\dots,I_{d_m}(h_m)) + \sqrt{t}Z_V\big)\left(\frac{1}{2\sqrt{t}}Z_V^i - \frac{1}{2\sqrt{1-t}}I_{d_i}(h_i)\right)\right].$$

By integrating by parts, we can write

$$\begin{aligned}
& E\left[\frac{\partial\varphi}{\partial x_i}\big(\sqrt{1-t}(I_{d_1}(h_1),\dots,I_{d_m}(h_m)) + \sqrt{t}Z_V\big)Z_V^i\right]\\
= \;& E\left\{E\left[\frac{\partial\varphi}{\partial x_i}\big(\sqrt{1-t}z + \sqrt{t}Z_V\big)Z_V^i\right]_{|z=(I_{d_1}(h_1),\dots,I_{d_m}(h_m))}\right\}\\
= \;& \sqrt{t}\sum_{j=1}^m V(i,j)\,E\left\{E\left[\frac{\partial^2\varphi}{\partial x_i\partial x_j}\big(\sqrt{1-t}z + \sqrt{t}Z_V\big)\right]_{|z=(I_{d_1}(h_1),\dots,I_{d_m}(h_m))}\right\}\\
= \;& \sqrt{t}\sum_{j=1}^m V(i,j)\,E\left[\frac{\partial^2\varphi}{\partial x_i\partial x_j}\big(\sqrt{1-t}(I_{d_1}(h_1),\dots,I_{d_m}(h_m)) + \sqrt{t}Z_V\big)\right].
\end{aligned}$$

By using e.g. (8.61) in order to perform the integration by parts, we can also write

$$\begin{aligned}
& E\left[\frac{\partial\varphi}{\partial x_i}\big(\sqrt{1-t}(I_{d_1}(h_1),\dots,I_{d_m}(h_m)) + \sqrt{t}Z_V\big)I_{d_i}(h_i)\right]\\
= \;& E\left\{E\left[\frac{\partial\varphi}{\partial x_i}\big(\sqrt{1-t}(I_{d_1}(h_1),\dots,I_{d_m}(h_m)) + \sqrt{t}z\big)I_{d_i}(h_i)\right]_{|z=Z_V}\right\}\\
= \;& \frac{\sqrt{1-t}}{d_i}\sum_{j=1}^m E\left\{E\left[\frac{\partial^2\varphi}{\partial x_i\partial x_j}\big(\sqrt{1-t}(I_{d_1}(h_1),\dots,I_{d_m}(h_m)) + \sqrt{t}z\big)\langle D[I_{d_i}(h_i)], D[I_{d_j}(h_j)]\rangle_{\mathfrak{H}}\right]_{|z=Z_V}\right\}\\
= \;& \frac{\sqrt{1-t}}{d_i}\sum_{j=1}^m E\left[\frac{\partial^2\varphi}{\partial x_i\partial x_j}\big(\sqrt{1-t}(I_{d_1}(h_1),\dots,I_{d_m}(h_m)) + \sqrt{t}Z_V\big)\langle D[I_{d_i}(h_i)], D[I_{d_j}(h_j)]\rangle_{\mathfrak{H}}\right].
\end{aligned}$$

Hence $\Psi'(t)$ equals

$$\frac{1}{2}\sum_{i,j=1}^{m} E\left[\frac{\partial^2\varphi}{\partial x_i\partial x_j}\big(\sqrt{1-t}(I_{d_1}(h_1),\ldots,I_{d_m}(h_m))+\sqrt{t}Z_V\big)\left(V(i,j)-\frac{1}{d_i}\langle D[I_{d_i}(h_i)],D[I_{d_j}(h_j)]\rangle_{\mathfrak{H}}\right)\right],$$

so that we get:

$$\begin{aligned}
\delta_2 &\leqslant \|\varphi''\|_\infty \sum_{i,j=1}^{m} E\left[\left|V(i,j)-\frac{1}{d_i}\langle D[I_{d_i}(h_i)],D[I_{d_j}(h_j)]\rangle_{\mathfrak{H}}\right|\right]\\
&\leqslant \|\varphi''\|_\infty \sum_{i,j=1}^{m} \sqrt{E\left[\left(V(i,j)-\frac{1}{d_i}\langle D[I_{d_i}(h_i)],D[I_{d_j}(h_j)]\rangle_{\mathfrak{H}}\right)^2\right]}\\
&= \|\varphi''\|_\infty \sum_{i,j=1}^{m} \frac{1}{d_i}\sqrt{\mathbf{Var}\big(\langle D[I_{d_i}(h_i)],D[I_{d_j}(h_j)]\rangle_{\mathfrak{H}}\big)}.
\end{aligned}$$

Step 4: Bounding $\mathbf{Var}\big(\langle D[I_{d_i}(h_i)],D[I_{d_j}(h_j)]\rangle_{\mathfrak{H}}\big)$. Assume, for instance, that $i\leqslant j$. We have

$$\begin{aligned}
&\langle D[I_{d_i}(h_i)],D[I_{d_j}(h_j)]\rangle_{\mathfrak{H}}\\
&= d_id_j\int_{\mathbb{R}} I_{d_i-1}\big(h_i(\cdot,a)\big)I_{d_j-1}\big(h_j(\cdot,a)\big)da\\
&= d_id_j\int_{\mathbb{R}}\sum_{r=0}^{d_i-1} r!\binom{d_i-1}{r}\binom{d_j-1}{r} I_{d_i+d_j-2-2r}\big(h_i(\cdot,a)\widetilde{\otimes}_r h_j(\cdot,a)\big)da \quad \text{by (2.28)}\\
&= d_id_j\int_{\mathbb{R}}\sum_{r=0}^{d_i-1} r!\binom{d_i-1}{r}\binom{d_j-1}{r} I_{d_i+d_j-2-2r}\big(h_i(\cdot,a)\otimes_r h_j(\cdot,a)\big)da\\
&= d_id_j\sum_{r=0}^{d_i-1} r!\binom{d_i-1}{r}\binom{d_j-1}{r} I_{d_i+d_j-2-2r}\Big(\int_{\mathbb{R}} h_i(\cdot,a)\otimes_r h_j(\cdot,a)da\Big)\\
&= d_id_j\sum_{r=0}^{d_i-1} r!\binom{d_i-1}{r}\binom{d_j-1}{r} I_{d_i+d_j-2-2r}\big(h_i\otimes_{r+1} h_j\big)\\
&= d_id_j\sum_{r=1}^{d_i} (r-1)!\binom{d_i-1}{r-1}\binom{d_j-1}{r-1} I_{d_i+d_j-2r}\big(h_i\otimes_r h_j\big)\\
&= d_id_j\sum_{r=1}^{d_i} (r-1)!\binom{d_i-1}{r-1}\binom{d_j-1}{r-1} I_{d_i+d_j-2r}\big(h_i\widetilde{\otimes}_r h_j\big).
\end{aligned}$$

Hence, if $d_i<d_j$, then $\mathbf{Var}\big(\langle D[I_{d_i}(h_i)],D[I_{d_j}(h_j)]\rangle_{\mathfrak{H}}\big)$ equals

$$d_i^2d_j^2\sum_{r=1}^{d_i}(r-1)!^2\binom{d_i-1}{r-1}^2\binom{d_j-1}{r-1}^2(d_i+d_j-2r)!\|h_i\widetilde{\otimes}_r h_j\|^2_{\mathfrak{H}^{\otimes(d_i+d_j-2r)}}$$

while, if $d_i = d_j$, it equals

$$d_i^4 \sum_{r=1}^{d_i-1} (r-1)!^2 \binom{d_i-1}{r-1}^4 (2d_i-2r)! \|h_i \widetilde{\otimes}_r h_j\|^2_{\mathfrak{H}^{\otimes(2d_i-2r)}}.$$

Now, let us stress the two following estimates. If $r < d_i \leqslant d_j$ then

$$\begin{aligned} \|h_i \widetilde{\otimes}_r h_j\|^2_{\mathfrak{H}^{\otimes(d_i+d_j-2r)}} &\leqslant \|h_i \otimes_r h_j\|^2_{\mathfrak{H}^{\otimes(d_i+d_j-2r)}} = \langle h_i \otimes_{d_i-r} h_i, h_j \otimes_{d_j-r} h_j \rangle_{\mathfrak{H}^{\otimes 2r}} \\ &\leqslant \|h_i \otimes_{d_i-r} h_i\|_{\mathfrak{H}^{\otimes 2r}} \|h_j \otimes_{d_j-r} h_j\|_{\mathfrak{H}^{\otimes 2r}} \\ &\leqslant \frac{1}{2} \left(\|h_i \otimes_{d_i-r} h_i\|^2_{\mathfrak{H}^{\otimes 2r}} + \|h_j \otimes_{d_j-r} h_j\|^2_{\mathfrak{H}^{\otimes 2r}} \right). \end{aligned}$$

If $r = d_i < d_j$, then

$$\|h_i \widetilde{\otimes}_{d_i} h_j\|^2_{\mathfrak{H}^{\otimes(d_j-d_i)}} \leqslant \|h_i \otimes_{d_i} h_j\|^2_{\mathfrak{H}^{\otimes(d_j-d_i)}} \leqslant \|h_i\|^2_{\mathfrak{H}^{\otimes d_i}} \|h_j \otimes_{d_j-d_i} h_j\|_{\mathfrak{H}^{\otimes 2d_i}}.$$

By putting all these estimates in the previous expression for $\mathbf{Var}\big(\langle D[I_{d_i}(h_i)], D[I_{d_j}(h_j)]\rangle_{\mathfrak{H}}\big)$, we get, using also (3.36), that

$$\frac{1}{d_i} \sqrt{\mathbf{Var}\big(\langle D[I_{d_i}(h_i)], D[I_{d_j}(h_j)]\rangle_{\mathfrak{H}}\big)} \leqslant \Delta_{ij},$$

for Δ_{ij} defined by (7.50).

This completes the proof of the theorem. ■

We now translate the bound in Theorem 7.2 into a bound for indicators of convex sets.

Corollary 7.3 *Let the notation and assumptions from Theorem 7.2 prevail. We consider the class $\mathcal{H}(\mathbb{R}^m)$ of indicator functions of measurable convex sets in $\mathbb{R}^m$. Let*

$$\begin{aligned} B_1 &= \frac{1}{2} \sum_{i=1}^m \Delta_{ii} + \sum_{1 \leqslant i < j \leqslant m} \Delta_{ij} \quad \textit{and} \\ B_2 &= C \left(\beta + \sqrt{\frac{8}{\pi}} \right) \left[\sum_{j=1}^m (16\sqrt{2}\beta)^{\frac{d_j-1}{3}} d_j! \right]^3 \sqrt{\max_{1 \leqslant j \leqslant m} \max_{1 \leqslant i \leqslant N_j} \mathrm{Inf}_i(f_j)}. \end{aligned}$$

1. *Assume that the covariance matrix V is the $m \times m$ identity matrix I_m. Then*

$$\sup_{h \in \mathcal{H}(\mathbb{R}^m)} \left| E\left[h\left(Q^1(\mathbf{X}), \ldots, Q^m(\mathbf{X}) \right) \right] - E[h(Z_V)] \right| \leqslant 8(B_1 + B_2)^{\frac{1}{4}} m^{\frac{3}{8}}.$$

2. *Assume that V is of rank $k \leqslant m$, and let $\Lambda = diag(\lambda_1, \ldots, \lambda_k)$ be the diagonal matrix with the non-zero eigenvalues of V on the diagonal. Let B be a $m \times k$ column*

orthonormal matrix (that is, $B^TB = I_k$ and $BB^T = I_m$), such that $V = B\Lambda B^T$, and let

$$b = \max_{i,j} \left(\Lambda^{-\frac{1}{2}} B^T\right)_{i,j}. \tag{7.52}$$

Then

$$\sup_{h\in\mathcal{H}(\mathbb{R}^m)} \left|E\left[h\left(Q^1(\mathbf{X}),\ldots,Q^m(\mathbf{X})\right)\right] - E[h(Z_V)]\right| \leqslant 8(b^2B_1 + b^3B_2)^{\frac{1}{4}} m^{\frac{3}{8}}.$$

Remark 7.4 1. We note that

$$\begin{aligned}&\sup_{z\in\mathbb{R}^m} \left|P\left[\left(Q^1(\mathbf{X}),\ldots,Q^m(\mathbf{X})\right) \leqslant z\right] - P[Z_V \leqslant z]\right| \\ \leqslant\ &\sup_{h\in\mathcal{H}(\mathbb{R}^m)} \left|E\left[h\left(Q^1(\mathbf{X}),\ldots,Q^m(\mathbf{X})\right)\right] - E[h(Z_V)]\right|.\end{aligned}$$

Thus Corollary 7.3 immediately gives a bound for Kolmogorov distance.

2. By using the bound for δ_2 derived in the proof of Theorem 7.2 above, and following the same line of reasoning as in the proof of Corollary 7.3, we have, by keeping the notation of Theorem 7.2, that if $\Delta_{ij} \to 0$ for all $i,j = 1,\ldots,m$ and $\max_{1\leqslant j\leqslant m} \max_{1\leqslant i\leqslant N_j} \mathrm{Inf}_i(f_j) \to 0$, then

$$\left(Q_{d_1}(N_1, f_1, \mathbf{G}),\ldots,Q_{d_m}(N_m, f_m, \mathbf{G})\right) \to \mathscr{N}_m(0, V), \quad \text{as } N_1,\ldots,N_j \to \infty,$$

in the Kolmogorov distance.

Proof. First assume that V is the identity matrix. We partially follow [37], and let Φ denote the standard normal distribution in $\mathbb{R}^m$, and ϕ the corresponding density function. For $h \in \mathcal{H}(\mathbb{R}^m)$, define the following smoothing:

$$h_t(x) = \int_{\mathbb{R}^m} h\left(\sqrt{t}y + \sqrt{1-t}x\right)\Phi(dy), \quad 0 < t < 1.$$

The key result, found for example in [10, Lemma 2.11], is that, for any probability measure Q on $\mathbb{R}^m$, for any $W \sim Q$ and $Z \sim \Phi$, and for any $0 < t < 1$, we have:

$$\sup_{h\in\mathcal{H}(\mathbb{R}^m)} |E[h(W)] - E[h(Z)]| \leqslant \frac{4}{3}\left[\sup_{h\in\mathcal{H}(\mathbb{R}^m)} |E[h_t(W)] - E[h_t(Z)]| + 2\sqrt{m}\sqrt{t}\right]. \tag{7.53}$$

Similarly as in [15, page 24], put

$$u(x,t,z) \;=\; (2\pi t)^{-\frac{m}{2}} \exp\left(-\sum_{i=1}^m \frac{(z_i - \sqrt{1-t}x_i)^2}{2t}\right),$$

and observe that $u(x,t,z)$ is the density function of the Gaussian vector $Y \sim \mathcal{N}(0, tI_m)$ taken in $z - \sqrt{1-t}x$. Substituting $z = \sqrt{t}y + \sqrt{1-t}x$, we get, for $0 < t < 1$,

$$h_t(x) = \int_{\mathbb{R}^m} h(z)u(x,t,z)dz.$$

By dominated convergence, we may differentiate under the integral and obtain, for the second derivatives,

$$\frac{\partial^2 h_t}{\partial x_i^2}(x) = -\frac{1-t}{t}\int_{\mathbb{R}^m} h(z)u(x,t,z)dz + \frac{1-t}{t^2}\int_{\mathbb{R}^m} h(z)(z_i - \sqrt{1-t}x_i)^2 u(x,t,z)dz$$

and, for $i \neq j$,

$$\frac{\partial^2 h_t}{\partial x_i \partial x_j}(x) = \frac{1-t}{t^2}\int_{\mathbb{R}^m} h(z)(z_i - \sqrt{1-t}x_i)(z_j - \sqrt{1-t}x_j)u(x,t,z)dz.$$

For the third partial derivatives,

$$\begin{aligned}\frac{\partial^3 h_t}{\partial x_i^3}(x) &= -3\frac{(1-t)^{\frac{3}{2}}}{t^2}\int_{\mathbb{R}^m} h(z)(z_i - \sqrt{1-t}x_i)u(x,t,z)dz \\ &\quad + \frac{(1-t)^{\frac{3}{2}}}{t^3}\int_{\mathbb{R}^m} h(z)(z_i - \sqrt{1-t}x_i)^3 u(x,t,z)dz\end{aligned}$$

and, for $i \neq j$,

$$\begin{aligned}\frac{\partial^3 h_t}{\partial x_i^2 \partial x_j}(x) &= -\frac{(1-t)^{\frac{3}{2}}}{t^2}\int_{\mathbb{R}^m} h(z)(z_j - \sqrt{1-t}x_j)u(x,t,z)dz \\ &\quad + \frac{(1-t)^{\frac{3}{2}}}{t^3}\int_{\mathbb{R}^m} h(z)(z_i - \sqrt{1-t}x_i)^2(z_j - \sqrt{1-t}x_j)u(x,t,z)dz,\end{aligned}$$

and, for i,j,k all distinct,

$$\begin{aligned}&\frac{\partial^3 h_t}{\partial x_i \partial x_j \partial x_k}(x) \\ &= \frac{(1-t)^{\frac{3}{2}}}{t^3}\int_{\mathbb{R}^m} h(z)(z_i - \sqrt{1-t}x_i)(z_j - \sqrt{1-t}x_j)(z_k - \sqrt{1-t}x_k)u(x,t,z)dz.\end{aligned}$$

Because $0 \leqslant h(z) \leqslant 1$ for all $z \in \mathbb{R}^m$, we may bound

$$\left|\frac{\partial^2 h_t}{\partial x_i^2}(x)\right| \leqslant \frac{1-t}{t} + \frac{1-t}{t^2}E[Y_i^2] = \frac{2(1-t)}{t}.$$

Similarly, for $i \neq j$,

$$\left|\frac{\partial^2 h_t}{\partial x_i \partial x_j}(x)\right| \leqslant \frac{1-t}{t^2}E[|Y_i|]E[|Y_j|] = \frac{2(1-t)}{\pi t}.$$

Thus, we have $\|h_t''\|_\infty \leqslant 1/t \leqslant 1/t^{3/2}$. Bounding the third derivatives in a similar fashion yields, for all i, j, k not necessarily distinct,

$$\begin{aligned}
&\left|\frac{\partial^3 h_t}{\partial x_i \partial x_j \partial x_k}(x)\right| \\
\leqslant\ & \frac{(1-t)^{\frac{3}{2}}}{t^3} \max\left\{3\, E[|Y_i|]\, t + E[|Y_i|^3]; E[|Y_j|]\, t + E[Y_i^2]E[|Y_j|]; E[|Y_i|]E[|Y_j|]E[|Y_k|]\right\},
\end{aligned}$$

and so $\|h_t'''\|_\infty \leqslant 1/t^{3/2}$. With (7.53) and Theorem 7.2, this gives that

$$\begin{aligned}
&\sup_{h\in\mathcal{H}(\mathbb{R}^m)} \left|E\left[h\left(Q^1(\mathbf{X}),\ldots,Q^m(\mathbf{X})\right)\right] - E[h(Z_V)]\right| \\
\leqslant\ & \frac{4}{3}\left[\sup_{h\in\mathcal{H}(\mathbb{R}^m)} \left|E\left[h_t\left(Q^1(\mathbf{X}),\ldots,Q^m(\mathbf{X})\right)\right] - E[h_t(Z_V)]\right| + 2\sqrt{m}\sqrt{t}\right] \\
\leqslant\ & \frac{8}{3}\sqrt{m}\sqrt{t} + \frac{4}{3}(B_1+B_2)t^{-\frac{3}{2}}.
\end{aligned}$$

This function is minimised for $t = \sqrt{\frac{3(B_1+B_2)}{2\sqrt{m}}}$, yielding that

$$\sup_{h\in\mathcal{H}(\mathbb{R}^m)} \left|E\left[h\left(Q^1(\mathbf{X}),\ldots,Q^m(\mathbf{X})\right)\right] - E[h(Z_V)]\right| \leqslant 8(B_1+B_2)^{\frac{1}{4}} m^{\frac{3}{8}},$$

as required.

For Point 2, write $W = (Q^1(\mathbf{X}),\ldots,Q^m(\mathbf{X}))$ for simplicity. For $h \in \mathcal{H}(\mathbb{R}^m)$, we have

$$E[h(W)] - E[h(Z_V)] = E[h(B\Lambda^{\frac{1}{2}} \times \Lambda^{-\frac{1}{2}}B^T W)] - E[h(B\Lambda^{\frac{1}{2}} \times \Lambda^{-\frac{1}{2}}B^T Z_V)].$$

Put $g(x) = h(B\Lambda^{\frac{1}{2}}x)$. Then, $g \in \mathcal{H}(\mathbb{R}^k)$ and, thanks to the inequality (7.53), we can write

$$\begin{aligned}
&\sup_{h\in\mathcal{H}(\mathbb{R}^m)} |E[h(W)] - E[h(Z_V)]| \\
\leqslant\ & \sup_{g\in\mathcal{H}(\mathbb{R}^k)} |E[g(\Lambda^{-\frac{1}{2}}B^T W)] - E[g(\Lambda^{-\frac{1}{2}}B^T Z_V)]| \\
\leqslant\ & \frac{4}{3}\left[\sup_{g\in\mathcal{H}(\mathbb{R}^k)} |E[g_t(\Lambda^{-\frac{1}{2}}B^T W)] - E[g_t(\Lambda^{-\frac{1}{2}}B^T Z_V)]| + 2\sqrt{k}\sqrt{t}\right].
\end{aligned}$$

We may bound the partial derivatives of $f_t(x) = g_t(\Lambda^{-\frac{1}{2}}B^T x)$ using the chain rule and (7.52), to give that

$$\|f_t''\|_\infty \leqslant b^2 t^{-\frac{3}{2}} \text{ and } \|f_t'''\|_\infty \leqslant b^3 t^{-\frac{3}{2}}.$$

Using Theorem 7.2 and minimising the bound in t as before gives the assertion; the only changes are that B_1 gets multiplied by b^2 and B_2 gets multiplied by b^3. ■

7.2 More universality

Here, we prove a slightly stronger version of Theorem 1.3 stated in Section 1.3. Precisely, we add the two conditions (2) and (3), making the criterion contained in Theorem 1.3 more effective for potential applications.

Theorem 7.5 *Let $\mathbf{G} = \{G_i : i \geqslant 1\}$ be a standard centered i.i.d. Gaussian sequence, and fix integers $m \geqslant 1$ and $d_1, ..., d_m \geqslant 2$. For every $j = 1, ..., m$, let $\{(N_n^{(j)}, f_n^{(j)}) : n \geqslant 1\}$ be a sequence such that $\{N_n^{(j)} : n \geqslant 1\}$ is a sequence of integers going to infinity, and each function $f_n^{(j)} : [N_n^{(j)}]^{d_j} \to \mathbb{R}$ is symmetric and vanishes on diagonals. Define $Q^j(n, \mathbf{G}) = Q_{d_j}(N_n^{(j)}, f_n^{(j)}, \mathbf{G})$, $n \geqslant 1$, according to (1.1). Assume that, for every $j = 1, ...m$, the sequence $E[Q^j(n, \mathbf{G})^2]$, $n \geqslant 1$, of variances is bounded. Let V be a $m \times m$ non-negative symmetric matrix, and let $\mathscr{N}_m(0, V)$ be a centered Gaussian vector with covariance matrix V. Then, as $n \to \infty$, the following two conditions are equivalent.*

1. *The vector $\{Q^j(n, \mathbf{G}) : j = 1, ..., m\}$ converges in law to $\mathscr{N}_m(0, V)$.*
2. *For all $i, j = 1, \ldots, m$, we have $E\big[Q^i(n, \mathbf{G})Q^j(n, \mathbf{G})\big] \to V(i, j)$ and $E\big[Q^i(n, \mathbf{G})^4\big] \to 3V(i, i)^2$ as $n \to \infty$.*
3. *For all $i, j = 1, \ldots, m$, we have $E\big[Q^i(n, \mathbf{G})Q^j(n, \mathbf{G})\big] \to V(i, j)$ and, for all $1 \leqslant i \leqslant m$ and $1 \leqslant r \leqslant d_i - 1$, we have $\|f_n^{(i)} \star_r f_n^{(i)}\|_{2d_i - 2r} \to 0$.*
4. *For every sequence $\mathbf{X} = \{X_i : i \geqslant 1\}$ of independent centered random variables, with unit variance and such that $\sup_i E|X_i|^3 < \infty$, the vector $\{Q^j(n, \mathbf{X}) : j = 1, ..., m\}$ converges in law to $\mathscr{N}_m(0, V)$ for the Kolmogorov distance.*

For the proof of Theorem 7.5, we need the following result, which consists in a collection of some of the findings contained in the papers by Peccati and Tudor [34]. Strictly speaking, the original statements contained in [34] only deal with positive definite covariance matrices: however, the extension to a non-negative matrix can be easily achieved by using the same arguments as in the "Step 3" of the proof of Theorem 7.2.

Theorem 7.6 *Fix integers $m \geqslant 1$ and $d_m \geqslant \ldots \geqslant d_1 \geqslant 1$. Let $V = \{V(i, j) : i, j = 1, ..., m\}$ be a $m \times m$ non-negative symmetric matrix. For any $n \geqslant 1$ and $i = 1, \ldots, m$, let $h_i^{(n)}$ belong to the d_i^{th} Gaussian chaos C_{d_i}. Assume that*

$$F^{(n)} = (F_1^{(n)}, \ldots, F_m^{(n)}) := (I_{d_1}(h_1^{(n)}), \ldots, I_{d_m}(h_m^{(n)})) \quad n \geqslant 1,$$

is such that

$$\lim_{n\to\infty} E[F_i^{(n)} F_j^{(n)}] = V(i, j), \quad 1 \leqslant i, j \leqslant m. \tag{7.54}$$

Then, as $n \to \infty$, the following four assertions are equivalent:

(i) *For every* $1 \leqslant i \leqslant m$, $F_i^{(n)}$ *converges in distribution to a centered Gaussian random variable with variance* $V(i,i)$.

(ii) *For every* $1 \leqslant i \leqslant m$, $E\left[(F_i^{(n)})^4\right] \to 3V(i,i)^2$.

(iii) *For every* $1 \leqslant i \leqslant m$ *and every* $1 \leqslant r \leqslant d_i - 1$, $\|h_i^{(n)} \otimes_r h_i^{(n)}\|_{\mathfrak{H}^{\otimes(2d_i-2r)}} \to 0$.

(iv) *The vector* $F^{(n)}$ *converges in distribution to the* d*-dimensional Gaussian vector* $\mathscr{N}_m(0,V)$.

Proof of Theorem 7.5. The equivalences (1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3) only consist in a reformulation of the previous Theorem 7.6, by taking into account the equality (3.36) and the fact that (since we suppose that the sequence $E[Q^j(n,\mathbf{G})^2]$ of variances is bounded, so that an hypercontractivity argument can be applied), if Point (1) is verified, then (7.54) holds. On the other hand, it is completely obvious that (4) implies (1), since $\mathbf{G}$ is a particular case of such an $\mathbf{X}$. So, it remains to prove the implication (1), (2), (3) $\Rightarrow$ (4). Let $Z_V = (Z_V^1, \ldots, Z_V^m) \sim \mathscr{N}_m(0,V)$. We have

$$\sup_{z\in\mathbb{R}^m} \left| P\left[Q^1(n,\mathbf{X}) \leqslant z_1, \ldots, Q^m(n,\mathbf{X}) \leqslant z_m\right] - P\left[Z_V^1 \leqslant z_1, \ldots, Z_V^m \leqslant z_m\right]\right| \leqslant \delta_n^{(a)} + \delta_n^{(b)}$$

with

$$\begin{aligned}
\delta_n^{(a)} &= \sup_{z\in\mathbb{R}^m} \big| P\left[Q^1(n,\mathbf{X}) \leqslant z_1, \ldots, Q^m(n,\mathbf{X}) \leqslant z_m\right] \\
&\qquad\qquad\qquad - P\left[Q^1(n,\mathbf{G}) \leqslant z_1, \ldots, Q^m(n,\mathbf{G}) \leqslant z_m\right]\big| \\
\delta_n^{(b)} &= \sup_{z\in\mathbb{R}^m} \left| P\left[Q^1(n,\mathbf{G}) \leqslant z_1, \ldots, Q^m(n,\mathbf{G}) \leqslant z_m\right] - P\left[Z_V^1 \leqslant z_1, \ldots, Z_V^m \leqslant z_m\right]\right|.
\end{aligned}$$

By assumption (3), we have that $\Delta_{ij} \to 0$ for all $i,j = 1, \ldots, m$ (with Δ_{ij} defined by (7.50)). Hence, Remark 7.4 (point 2) implies that $\delta_n^{(b)} \to 0$. By assumption (3) (for $r = d_j - 1$) and (1.14)-(1.15), we get that $\max_{1\leqslant i\leqslant N_n^{(j)}} \mathrm{Inf}_i(f_n^{(j)}) \to 0$ as $n \to \infty$ for all $j = 1, \ldots, m$. Hence, Corollary 7.3 implies that $\delta_n^{(a)} \to 0$, which completes the proof.

■

8 Some proofs based on Malliavin calculus and Stein's method

8.1 The language of Malliavin calculus

Let $\mathbf{G} = \{G_i : i \geqslant 1\}$ be an i.i.d. sequence of Gaussian random variables with zero mean and unit variance. In what follows, concerning the Wiener Gaussian chaos, we will systematically use the definitions and notation that have been introduced in Section 2. In particular, we shall encode the structure of random variables belonging to some Wiener chaos by means of increasing (tensor) powers of a fixed real separable Hilbert space $\mathfrak{H}$. We recall that the first Wiener chaos of $\mathbf{G}$ (which coincides with the Gaussian space generated by $\mathbf{G}$) is the L^2-closed Hilbert space composed of random variables of the type $I_1(h)$, where $h \in \mathfrak{H}$. We shall denote by $L^2(\mathbf{G})$ the space of all $\mathbb{R}$-valued random elements F that are measurable with respect to $\sigma\{\mathbf{G}\}$ and verify $E[F^2] < \infty$. Also, $L^2(\Omega; \mathfrak{H})$ denotes the space of all $\mathfrak{H}$-valued random elements u, that are measurable with respect to $\sigma\{\mathbf{G}\}$ and verify the relation $E\big[\|u\|^2_{\mathfrak{H}}\big] < \infty$.

We stress (see again [27, Ch. 1] or [11]) that any random variable F belonging to $L^2(\mathbf{G})$ admits the following chaotic expansion:

$$F = E[F] + \sum_{d=1}^{\infty} I_d(h_d), \tag{8.55}$$

where the series converges in $L^2(\mathbf{G})$ and the symmetric kernels $h_d \in \mathfrak{H}^{\odot d}$, $d \geqslant 1$, are uniquely determined by F. For convenience we also put

$$I_0(f_0) = E[F],$$

so that we can write

$$F = \sum_{d=0}^{\infty} I_d(h_d).$$

In the particular case where $\mathfrak{H} = L^2(A, \mathscr{A}, \mu)$, with $(A, \mathscr{A})$ a measurable space and μ a σ-finite and non-atomic measure, $\mathfrak{H}^{\odot d} = L^2_s(A^d, \mathscr{A}^{\otimes d}, \mu^{\otimes d})$ is the space of symmetric and square integrable functions on A^d.

Let $\mathscr{S}$ be the set of all smooth cylindrical random variables of the form

$$F = g\big(I_1(\phi_1), \dots, I_1(\phi_n)\big)$$

where $n \geqslant 1$, $g : \mathbb{R}^n \to \mathbb{R}$ is a smooth function with compact support and $\phi_i \in \mathfrak{H}$. The Malliavin derivative of F (with respect to $\mathbf{G}$) is the element of $L^2(\Omega; \mathfrak{H})$ defined as

$$DF \;=\; \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}\big(I_1(\phi_1), \dots, I_1(\phi_n)\big)\phi_i.$$

By iteration, one can define the mth derivative $D^m F$ (which is an element of $\mathrm{L}^2(\Omega, \mathfrak{H}^{\otimes m})$) for every $m \geqslant 2$. For $m \geqslant 1$, $\mathbb{D}^{m,2}$ denotes the closure of $\mathscr{S}$ with respect to the norm $\|\cdot\|_{m,2}$, defined by the relation

$$\|F\|^2_{m,2} = E[F^2] + \sum_{i=1}^{m} E\big[\|D^i F\|^2_{\mathfrak{H}^{\otimes i}}\big].$$

We have $DI_1(h) = h$ for every $h \in \mathfrak{H}$. The Malliavin derivative D satisfies the following *chain rule*: if $g : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable and has bounded partial derivatives, and if $(F_1, ..., F_n)$ is a vector of elements of $\mathbb{D}^{1,2}$, then $g(F_1, \dots, F_n) \in \mathbb{D}^{1,2}$ and

$$Dg(F_1,\dots,F_n) = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(F_1,\dots,F_n)DF_i. \tag{8.56}$$

A careful application of the multiplication formula (2.28) shows that (8.56) continues to hold when $(F_1, ..., F_n)$ is a vector of multiple integrals (of possibly different orders) and g is a polynomial in n variables. We denote by δ the adjoint of the operator D, also called the *divergence operator*. A random element $u \in L^2(\Omega;\mathfrak{H})$ belongs to the domain of δ, noted Domδ, if and only if it satisfies

$$|E\langle DF, u\rangle_{\mathfrak{H}}| \leqslant c_u \sqrt{E[F^2]} \quad \text{for any } F \in \mathscr{S},$$

for some constant c_u depending only on u. If $u \in \mathrm{Dom}\delta$, then the random variable $\delta(u)$ is defined by the duality relationship (customarily called "integration by parts formula"):

$$E(F\delta(u)) = E\langle DF, u\rangle_{\mathfrak{H}}, \tag{8.57}$$

which holds for every $F \in \mathbb{D}^{1,2}$. If $\mathfrak{H} = L^2(A, \mathscr{A}, \mu)$ (with μ non-atomic), then the derivative of a random variable $F = E(F) + \sum_{d\geqslant 1} I_d(h_d)$ can be identified with the element of $L^2(A\times\Omega)$ given by

$$D_a F = \sum_{d=1}^{\infty} dI_{d-1}\big(h_d(\cdot, a)\big), \quad a \in A. \tag{8.58}$$

Next we define the operator L, called the *infinitesimal generator of the Ornstein-Uhlenbeck semigroup*. It acts on square-integrable random variables as follows: a random variable $F = E(F) + \sum_{d\geqslant 1} I_d(h_d)$ is in the domain of L, noted Dom$L$, if and only if $\sum_{d\geqslant 1} dI_d(h_d)$ is convergent in L^2, and in this case $LF = -\sum_{d\geqslant 1} dI_d(h_d)$. The operator L verifies the following crucial properties: (i) DomL $= \mathbb{D}^{2,2}$, and (ii) a random variable F is in Dom$L$ if and only if $F \in \mathrm{Dom}\delta D$ (i.e. $F \in \mathbb{D}^{1,2}$ and $DF \in \mathrm{Dom}\delta$), and in this case one has that $\delta(DF) = -LF$. We also define the operator L^{-1}, which is the *pseudo-inverse* of L, as follows: for every centered random variable $F = \sum_{d\geqslant 1} I_d(h_d)$ of $L^2(\mathbf{G})$, we set $L^{-1}F = \sum_{d\geqslant 1} -\frac{1}{d} I_d(h_d)$. Note that L^{-1} is an operator with values in $\mathbb{D}^{2,2}$.

Now consider a centered $F \in L^2(\mathbf{G})$. By using the relations $F = L(L^{-1}F) = \delta(-DL^{-1}F)$, one deduces from (8.57) and (8.56) that, for every $G \in \mathbb{D}^{1,2}$ and for every continuously differentiable $g : \mathbb{R} \to \mathbb{R}$ with a bounded derivative, we have

$$E[g(G)F] = E[g'(G)\langle DG, -DL^{-1}F\rangle_{\mathfrak{H}}]. \tag{8.59}$$

In particular, with $g(x) = x$, we get

$$E[GF] = E[\langle DG, -DL^{-1}F\rangle_{\mathfrak{H}}]. \tag{8.60}$$

Moreover, if $F = I_d(h)$, with $h \in \mathfrak{H}^{\odot d}$, then $L^{-1}F = -\frac{1}{d}F$, and therefore (8.59) and (8.60) become

$$E[g(G)F] = \frac{1}{d}E[g(G)\langle DG, DF\rangle_{\mathfrak{H}}] \quad \text{and} \quad E[GF] = \frac{1}{d}E[\langle DG, DF\rangle_{\mathfrak{H}}]. \tag{8.61}$$

Let $h \in \mathfrak{H}^{\odot d}$ with $d \geqslant 2$, and let $s \geqslant 0$ be an integer. The following identity is obtained by taking $F = I_d(h)$ and $G = F^{s+1}$ in formula (8.61), and then by applying (8.56):

$$E\big[I_d(h)^{s+2}\big] = \frac{s+1}{d} E\big[I_d(h)^s \|DI_d(h)\|^2_{\mathfrak{H}}\big]. \tag{8.62}$$

8.2 Relations following from Stein's method

Originally introduced in [41, 42], *Stein's method* can be described as a collection of probabilistic techniques, allowing to compute explicit bounds on the distance between the laws of random variables by means of differential operators. The reader is referred to [2] and [36], and the references therein, for an introduction to these techniques. The following statement contains four bounds which can be obtained by means of a combination of Malliavin calculus and Stein's method. Points 1, 2 and 4 have been proved in [22], whereas the content of Point 3 is new. Our proof of such a bound gives an explicit example of the interaction between Stein's method and Malliavin calculus. We also introduce the following notation: for every centered $F \in \mathbb{D}^{1,2}$, we write

$$T_0(F) = \sqrt{\mathbf{Var}(\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}})}, \tag{8.63}$$

where **Var** indicates variance (observe that, due to (8.60), $E(\langle DF, -DL^{-1}F\rangle_{\mathfrak{H}}) = E(F^2)$). Note that,

$$\text{if } F = I_d(f), \text{ then } T_0(F) = \sqrt{\mathbf{Var}\left(\frac{1}{d}\|DF\|^2_{\mathfrak{H}}\right)}. \tag{8.64}$$

Proposition 8.1 *Consider $F = I_d(h)$ with $d \geqslant 1$ and $h \in \mathfrak{H}^{\odot d}$, and let Z and Z_ν have respectively a $\mathscr{N}(0,1)$ and a $\chi^2(\nu)$ distribution ($\nu \geqslant 1$). If $E(F^2) = 1$ then the following three bounds hold:*

1. $d_{TV}(F, Z) \leqslant 2T_0(F)$.
2. $d_W(F, Z) \leqslant T_0(F)$.
3. *For every thrice differentiable function $\varphi : \mathbb{R} \to \mathbb{R}$ such that $\|\varphi'''\| < \infty$,*

$$|E[\varphi(F)] - E[\varphi(Z)]| \leqslant C_* \times T_0(F),$$

where C_* is given in (3.34).

If $E(F^2) = 2\nu$ then the following bound is in order:

4. $d_{BW}(F, Z_\nu) \leqslant \max\left\{\sqrt{\frac{2\pi}{\nu}}, \frac{1}{\nu} + \frac{2}{\nu^2}\right\} \sqrt{E[(2\nu + 2F - \langle DF, -DL^{-1}F\rangle_{\mathfrak{H}})^2]}$.

Proof. Point 1 and 2 are proved in [22, Theorem 3.1], while Point 4 is proved in [22, Theorem 3.11]. To prove Point 3, fix φ as in the statement, and consider the *Stein equation*

$$f'(x) - xf(x) = \varphi(x) - E[\varphi(Z)], \quad x \in \mathbb{R}. \tag{8.65}$$

It is easily seen that a solution of (8.65) is given by

$$f_\varphi(x) = e^{x^2/2}\int_{-\infty}^{x} (\varphi(y) - E[\varphi(Z)])e^{-y^2/2}dy. \tag{8.66}$$

According to the forthcoming Lemma 8.2, the following bound is in order:

$$|f_\varphi'(x)| \leqslant K_*(1 + |x| + |x|^2 + |x|^3), \tag{8.67}$$

where K_* is given in (3.33). Now use (8.61) with $g = f_\varphi$ and $G = F$, as well as a standard aproximation argument to take into account that f_φ' is not necessarily bounded, in order to write

$$\begin{aligned}
&|E[\varphi(F)] - E[\varphi(Z)]| = |E[f_\varphi'(F) - Ff_\varphi(F)]| \\
&= \left| E\left[f_\varphi'(F)\left(1 - \frac{1}{d}\|DF\|_{\mathfrak{H}}^2\right)\right]\right| \\
&\leqslant K_* E\left[\left(1 + |F| + |F|^2 + |F|^3\right)\left|1 - \frac{1}{d}\|DF\|_{\mathfrak{H}}^2\right|\right].
\end{aligned}$$

By applying the Cauchy-Schwarz inequality four times, and by exploiting Proposition 2.7, one infers

$$\begin{aligned}
& K_* E\left[\left(1 + |F| + |F|^2 + |F|^3\right)\left|1 - \frac{1}{d}\|DF\|_{\mathfrak{H}}^2\right|\right] \\
\leqslant\ & C_*\sqrt{E\left[\left(1 - \frac{1}{d}\|DF\|_{\mathfrak{H}}^2\right)^2\right]} = C_* T_0(F),
\end{aligned}$$

which is the desired conclusion. ■

Lemma 8.2 *The function f_φ defined in (8.66) verifies relation (8.67).*

Proof. We denote by Z a Gaussian random variable with zero mean and unit variance. Observe that

$$E\,|Z|^3 = \frac{2\sqrt{2}}{\sqrt{\pi}} \quad \text{and} \quad E\,|Z| = \frac{\sqrt{2}}{\sqrt{\pi}}. \tag{8.68}$$

We want to bound the quantity $\left|f_\varphi'(x)\right|$, where: (i) φ is such that $\varphi(x) = \varphi(0) + \varphi'(0)\,x + \varphi''(0)\,x^2/2 + R(x)$, with $|R(x)| \leqslant \|\varphi'''\|_\infty |x|^3/6$, and (ii) f_φ is given by (8.66), and consequently

$$f_\varphi'(x) = \varphi(x) - E\varphi(Z) + xf_\varphi(x) = A(x) + B(x), \tag{8.69}$$

with $A(x) := \varphi(x) - E\varphi(Z)$ and $B(x) := xf_\varphi(x)$. It will become clear later on that our bounds on $\left|f'_\varphi(x)\right|$ do not depend on the sign of x, so that in what follows we will uniquely focus on the case $x > 0$. Due to the assumptions on φ, we have that

$$\begin{aligned} A(x) &= \varphi'(0)\,x + \frac{1}{2}\varphi''(0)\,x^2 + R(x) + C \\ &:= ax + bx^2 + R(x) + C, \quad \text{where} \\ -C &= \frac{\varphi''(0)}{2} + ER(Z), \end{aligned}$$

(note that the term $\varphi(0)$ simplifies) and also, by using (8.68),

$$|C| \leqslant \frac{|\varphi''(0)|}{2} + \frac{\|\varphi'''\|_\infty}{3}\frac{\sqrt{2}}{\sqrt{\pi}} := C'$$

and (recall that $x > 0$)

$$\begin{aligned} |A(x)| &\leqslant |\varphi'(0)|\,x + \frac{1}{2}|\varphi''(0)|\,x^2 + \frac{1}{6}\|\varphi'''\|_\infty x^3 + C' \\ &= |a|\,x + |b|\,x^2 + \gamma x^3 + C' \\ \gamma &:= \frac{1}{6}\|\varphi'''\|_\infty . \end{aligned}$$

On the other hand, since $EA(Z) = 0$ by construction,

$$\begin{aligned} |B(x)| &= xe^{\frac{x^2}{2}}\left|\int_x^{+\infty} A(y)\,e^{-\frac{y^2}{2}}dy\right| \\ &\leqslant xe^{\frac{x^2}{2}}\int_x^{+\infty}\left[C' + |a|\,y + |b|\,y^2 + \gamma y^3\right]e^{-\frac{y^2}{2}}dy \\ &:= Y_1(x) + Y_2(x) + Y_3(x) + Y_4(x). \end{aligned}$$

We now evaluate the four terms Y_i separately (observe that each of them is positive):

$$\begin{aligned} Y_1(x) &= C'xe^{\frac{x^2}{2}}\int_x^{+\infty} e^{-\frac{y^2}{2}}dy \leqslant C'e^{\frac{x^2}{2}}\int_x^{+\infty} ye^{-\frac{y^2}{2}}dy = C'; \\ Y_2(x) &= xe^{\frac{x^2}{2}}\int_x^{+\infty}|a|\,ye^{-\frac{y^2}{2}}dy = |a|\,x; \\ Y_3(x) &= xe^{\frac{x^2}{2}}\int_x^{+\infty}|b|\,y^2e^{-\frac{y^2}{2}}dy \leqslant e^{\frac{x^2}{2}}\int_x^{+\infty}|b|\,y^3e^{-\frac{y^2}{2}}dy \\ &= |b|\left(x^2+2\right) = 2\,|b| + |b|\,x^2; \\ Y_4(x) &= xe^{\frac{x^2}{2}}\int_x^{+\infty}\gamma y^3e^{-\frac{y^2}{2}}dy = \gamma x\left(x^2+2\right) = \gamma x^3 + 2\gamma x. \end{aligned}$$

By combining the above bounds with (8.69), one infers that

$$\begin{aligned} \left|f'_\varphi(x)\right| &\leqslant 2C' + 2\,|b| + x\,(2\,|a| + 2\gamma) + x^2 2\,|b| + x^3 2\gamma \\ &\leqslant \max\{2C' + 2\,|b|\,; 2\,|a| + 2\gamma; 2\,|b|\,; 2\gamma\} \times \left(1 + x + x^2 + x^3\right) \\ &= \max\{2C' + 2\,|b|\,; 2\,|a| + 2\gamma\} \times \left(1 + x + x^2 + x^3\right), \end{aligned}$$

which yields the desired conclusion. ■

8.3 Proof of Theorem 3.1

Let $F = I_d(h)$, $h \in \mathfrak{H}^{\odot d}$. In view of Proposition 8.1, it is sufficient to show that

$$T_0(F) = T_1(F) \leqslant T_2(F).$$

Relation (3.42) in [22] yields that

$$\frac{1}{d}\|DF\|^2_{\mathfrak{H}} = E(F^2) + d\sum_{r=1}^{d-1}(r-1)!\binom{d-1}{r-1}^2 I_{2d-2r}\left(h\widetilde{\otimes}_r h\right), \tag{8.70}$$

which, by taking the orthogonality of multiple integrals of different orders into account, yields

$$\mathbf{Var}\left(\frac{1}{d}\|DF\|^2_{\mathfrak{H}}\right) = d^2\sum_{r=1}^{d-1}(r-1)!^2\binom{d-1}{r-1}^4 (2d-2r)!\|h\widetilde{\otimes}_r h\|^2_{\mathfrak{H}^{\otimes 2(d-r)}}, \tag{8.71}$$

and consequently $T_0(F) = T_1(F)$. From the multiplication formula (2.28)) we get

$$F^2 = \sum_{r=0}^{d} r!\binom{d}{r}^2 I_{2d-2r}(h\widetilde{\otimes}_r h). \tag{8.72}$$

To conclude the proof, we use (8.62) with $s = 2$, combined with (8.70) and (8.72), as well as the assumption that $E(F^2) = 1$, to get that

$$\begin{aligned} E[F^4] - 3 &= \frac{3}{d}E(F^2\|DF\|^2_{\mathfrak{H}}) - 3(d!\|h\|^2_{\mathfrak{H}^{\otimes d}})^2 \\ &= 3d\sum_{r=1}^{d-1} r!\,(r-1)!\binom{d}{r}^2\binom{d-1}{r-1}^2 (2d-2r)!\|h\widetilde{\otimes}_r h\|^2_{\mathfrak{H}^{\otimes 2(d-r)}}. \end{aligned}$$

Hence

$$\mathbf{Var}\left(\frac{1}{d}\|DF\|^2_{\mathfrak{H}}\right) \leqslant \frac{d-1}{3d}[E(F^4) - 3],$$

thus yielding the desired inequality $T_1(F) \leqslant T_2(F)$.

8.4 Proof of Theorem 3.6

Let $F = I_d(h)$, $h \in \mathfrak{H}^{\odot d}$. In view of Proposition 8.1 and since $L^{-1}F = -\frac{1}{d}F$, it is sufficient to show that

$$\sqrt{E\left[\left(2\nu + 2F - \frac{1}{d}\|DF\|^2_{\mathfrak{H}}\right)^2\right]} = T_3(F) \leqslant T_4(F).$$

By taking into account the orthogonality of multiple integrals of different orders, relation (8.70) yields

$$
\begin{aligned}
E\left[\left(2\nu+2F-\frac{1}{d}\|DF\|_{\mathfrak{H}}^2\right)^2\right] &= 4d!\left\|h-\frac{d!^2}{4(d/2)!^3}h\widetilde{\otimes}_{d/2}h\right\|_{\mathfrak{H}^{\otimes d}}^2 \\
&+ d^2\sum_{\substack{r=1,\ldots,d-1\\ r\neq d/2}}(r-1)!^2\binom{d-1}{r-1}^4(2d-2r)!\left\|h\widetilde{\otimes}_r h\right\|_{\mathfrak{H}^{\otimes(2d-2r)}}^2,
\end{aligned}
$$

and consequently $T_3(F)=\sqrt{E\left[\left(2\nu+2F-\frac{1}{d}\|DF\|_{\mathfrak{H}}^2\right)^2\right]}$. On the other hand, by combining (8.62) (for $s=1$ and $s=2$) with (8.72), we get, still by taking into account the orthogonality of multiple integrals of different orders:

$$
\begin{aligned}
E[F^4]-12E[F^3] &= 12\nu^2-48\nu+24d!\left\|h-\frac{d!^2}{4(d/2)!^3}h\widetilde{\otimes}_{d/2}h\right\|_{\mathfrak{H}^{\otimes d}}^2 \\
&+ 3d\sum_{\substack{r=1,\ldots,d-1\\ r\neq d/2}} r!\,(r-1)!\binom{d}{r}^2\binom{d-1}{r-1}^2(2d-2r)!\left\|h\widetilde{\otimes}_r h\right\|_{\mathfrak{H}^{\otimes(2d-2r)}}^2.
\end{aligned}
$$

It is now immediate to deduce that $T_3(F)\leqslant T_4(F)$.

Acknowledgement. Part of this paper has been written while the three authors were visiting the Institute for Mathematical Sciences of the National University of Singapore, in the occasion of the program "Progress in Stein's method" (January 5 - February 6 2009). We heartily thank Andrew Barbour, Louis Chen and Kwok Pui Choi for their kind hospitality and generous support.

References

[1] R. Blei (2001). *Analysis in Integer and Fractional Dimensions*. Cambridge University Press.

[2] L.H.Y. Chen and Q.-M. Shao (2004). Normal approximation under local dependence. *Ann. Probab.*32, 1985-2028.

[3] L.H.Y. Chen and Q.-M. Shao (2005). Stein's method for normal approximation. In: *An introduction to Stein's method*, 1-59. Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap. **4**, Singapore Univ. Press, Singapore, 2005.

[4] Y. Davidov and V. Rotar' (2007). On asymptotic proximity of distributions. Preprint.

[5] P. de Jong (1989). Central Limit Theorems for generalized multilinear forms. Tract **61**. CWI, Amsterdam.

[6] P. de Jong (1990). A central limit theorem for generalized multilinear forms. *J. Mult. Anal.* **34**, 275-289.

[7] W. Feller (1965). *An Introduction to Probability Theory and its Applications, Vol. II.* Second Edition. Wiley, New York.

[8] A.L. Gibbs and F. Su (2002). On choosing and bounding probability metrics. *International Statistical Review* **70**, 419-435.

[9] F. Götze and A.N. Tikhomirov (2002). Asymptotic distribution of quadratic forms and applications. *J. Theoret. Probab.* **15**(2), 423-475.

[10] F. Götze (1991). On the rate of convergence in the multivariate CLT. *Ann. Probab.* **19**, 724-739.

[11] S. Janson (1997). *Gaussian Hilbert Spaces.* Cambridge University Press.

[12] O. Kallenberg (1997). *Foundations of modern probability.* Springer.

[13] M. Ledoux and M. Talagrand (1991). *Probability on Banach spaces.* Springer-Verlag, Berlin Heidelberg New York.

[14] J.W. Lindeberg (1922). Eine neue Herleitung des exponential-Gesetzes in der Warscheinlichkeitsrechnung. *Math. Zeitschrift.* **15**, 211-235

[15] W. L. Loh (2007). A multivariate central limit theorem for randomised orthogonal array sampling designs in computer experiments. *Manuscript*, available at `http://arxiv.org/abs/0708.0656v1`.

[16] P. Malliavin (1997). *Stochastic Analysis.* Springer-Verlag, Berlin, Heidelberg, New York.

[17] K.V. Mardia, J.T. Kent, and J.M. Bibby (1979). *Multivariate Analysis.* Academic Press, London, San Diego, New York, Boston, Sydney, Tokyo, Toronto.

[18] P.-A. Meyer (1993). *Quantum Probability for Probabilists.* LNM 1538. Springer-Verlag, Berlin, Heidelberg, New York.

[19] E. Mossel (2008). Gaussian bounds for noise correlation of functions. To appear in: *GAFA.*

[20] E. Mossel, R. O'Donnell and K. Oleszkiewicz (2008). Noise stability of functions with low influences: variance and optimality. To appear in: *Ann. Math.*.

[21] I. Nourdin and G. Peccati (2009a). Non-central convergence of multiple integrals. To appear in: *Ann. Probab.*

[22] I. Nourdin and G. Peccati (2009b). Stein's method on Wiener chaos. To appear in: *Probab. Theory Rel. Fields.*

[23] I. Nourdin and G. Peccati (2009c). Stein's method and exact Berry-Essen bounds for functionals of Gaussian fields. To appear in: *Ann. Probab.*.

[24] I. Nourdin, G. Peccati and A. Réveillac (2009). Multivariate normal approximation using Stein's method and Malliavin calculus. To appear in: *Ann. Inst. H. Poincaré Probab. Statist.*.

[25] I. Nourdin, G. Peccati and G. Reinert (2008). Stein's method and stochastic analysis of Rademacher functionals. Preprint.

[26] I. Nourdin, G. Peccati and G. Reinert (2009). Second order Poincaré inequalities and CLTs on Wiener space. To appear in: *J. Funct. Anal.*

[27] D. Nualart (2006). *The Malliavin calculus and related topics.* Springer-Verlag, Berlin, 2nd edition.

[28] D. Nualart and S. Ortiz-Latorre (2008). Central limit theorems for multiple stochastic integrals and Malliavin calculus. *Stochastic Process. Appl.* **118** (4), 614-628.

[29] D. Nualart and G. Peccati (2005). Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.* **33** (1), 177-193.

[30] N. Obata (1994). *White noise calculus and Fock space.* Lecture Notes in Mathematics **1577**. Springer-Verlag, Berlin.

[31] Ogura, H. (1972). Orthogonal functionals of the Poisson process. *IEEE Trans. Inform. Theory IT*, **18**, 473-481.

[32] G. Peccati, J.-L. Solé, M.S. Taqqu and F. Utzet (2008). Stein's method and normal approximation of Poisson functionals. To appear in: *Ann. Probab.*.

[33] G. Peccati and M.S. Taqqu (2008). Moments, cumulants and diagram formulae for non-linear functionals of random measure (Survey). Preprint.

[34] Peccati, G. and Tudor, C.A. (2004). Gaussian limits for vector-valued multiple stochastic integrals. In: *Séminaire de Probabilités XXXVIII*, 247-262, Springer Verlag

[35] N. Privault (1994). Chaotic and variational calculus in discrete and continuous time for the Poisson process. *Stochastics and Stochastics Reports* **51**, 83-109.

[36] G. Reinert (2005). Three general approaches to Stein's method. In: *An introduction to Stein's method*, 183-221. Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap. **4**, Singapore Univ. Press, Singapore.

[37] Y. Rinott and V. Rotar (1996). A multivariate CLT for local dependence with $n^{-1/2}\log n$ rate and applications to multivariate graph related statistics. *J. Multivariate Anal.* **56**, 333–350.

[38] G.-C. Rota and C. Wallstrom (1997). Stochastic integrals: a combinatorial approach. *Ann. Probab.* **25**(3), 1257-1283.

[39] V. I. Rotar' (1975). Limit theorems for multilinear forms and quasipolynomial functions. *Teor. Verojatnost. i Primenen.*, **20**(3), 527546.

[40] V. I. Rotar' (1979). Limit theorems for polylinear forms. *J. Multivariate Anal.* **9**, 511-530.

[41] Ch. Stein (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In: *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. II: Probability theory*, 583-602. Univ. California Press, Berkeley, CA.

[42] Ch. Stein (1986). *Approximate computation of expectations.* Institute of Mathematical Statistics Lecture Notes - Monograph Series, **7**. Institute of Mathematical Statistics, Hayward, CA.

[43] D. Surgailis (2003a). CLTs for Polynomials of Linear Sequences: Diagram Formulae with Applications. In: *Long Range Dependence*. Birkhäuser, Basel, 111-128.

[44] M. Talagrand (2003). Spin Glasses: A Challenge for Mathematicians. Cavity and Mean Fields Models. Springer, Berlin.

[45] T. Tao and V. Vu (2008). Random matrices: Universality of ESDs and the circular law. Preprint.

[46] H.F. Trotter (1959). An elementary proof of the central limit theorem. *Arch. Math.* **10**, 226-234.